

Playing the trump card of uncertainty

Republicans in the United States are promoting legislation that would insist on scientific risk analysis before any tightening of environmental regulation. The proposals are not acceptable.

PUBLICATION last month of studies showing falling sperm counts in British men is a reminder of the insidious threat posed by environmental pollutants. Not that pollutants are necessarily the source of the falling sperm counts reported over recent years. Changes in women's lifestyles may also have affected the development of male fetuses, while men's smoking and prolonged driving in tight trousers have also been suggested as possible causes. The statistics of sperm decline are themselves controversial. But endocrine-disrupting chemicals affecting fetal testes are prime suspects.

In assessing the potential impact of such pollutants, the current limitations of science are all too apparent. When, as with trace chemicals and particulates, low doses may be important, epidemiology becomes important. But it can accomplish only so much. True, it can provide useful information at lower levels of exposure than can be explored in laboratory tests on animals; it can also help to relate the results of such tests to effects on humans. But epidemiologists acknowledge that their studies of human populations can be unavoidably afflicted by statistical uncertainty, because samples sometimes have to be small. Such factors become important when apparently small consequences of low exposure may amplify the effects of common diseases. And although progress is being made in understanding mechanisms by which, for example, dioxins may harm humans, the number of possible molecular pathways in such cases is still considerable, while mechanisms are even less well understood for atmospheric particles that are suspected carcinogens.

Where the issues of pollution have become politicized, uncertainties within the science provide a notorious opportunity for the conflicting parties in any environmental dispute. That was the case when the chemical industry used its knowledge of atmospheric models to resist restrictions on chlorofluorocarbons, and when Greenpeace and UK Nirex, responsible in the United Kingdom for nuclear waste disposal, use hydrogeological models in disputes over Nirex's plans to test a site for a subterranean repository.

This week's Briefing (pages 10–14) focuses on the role of science within two debates which have become even more highly charged: the unprecedented involvement of right-wing politicians in environmental legislation in the United States, exemplified by debates over atmospheric particles; and the extraordinary (but possibly transitory) victory of Greenpeace over the Shell oil company in an environmentally innocuous plan (as it appeared to its originators) to sink an oil storage buoy — the Brent Spar — in deep water off northwest Scotland.

The common theme is science-based risk analysis. This can be seen pessimistically as a way of highlighting scientific uncertainties. After all, a classical risk analysis derives the probabilities of various outcomes by analysing all the possible paths by which each may materialize. So, for example, in assessing risks to human health from chemically contaminated fish, the US Environmental Protection Agency has presented a framework in which, for just one chemical, estimates of many factors ranging from possible environmental pathways to coefficients of absorption — all uncertain in conceptual terms or in measurement terms or both — have to be amalgamated to make a quantitative risk assessment possible. The accumulation of uncertainties looks daunting, and provides ample opportunity for groups with vested interests to promote a biased perspective. But such risk analysis, by assessing

scenarios comprehensively, also has valuable potential for establishing common ground.

Suppose, in the case of Brent Spar, that the environmental impact analyses made by Shell are vindicated by the working group set up by the UK government. Greenpeace's accusations that Shell's science was deficient would be seen to be exaggerated, and the common ground between the factions might have increased. But Greenpeace is unlikely to change its position as a result because its overriding principle of environmental precaution in the face of residual scientific uncertainties and future dumping possibilities would still be politically tenable.

In the United States, Republican zealots and some industries are likely to go on playing the same card of uncertainty and precaution in the opposite direction for as long as they can, while — as the Briefing describes — they propose risk-analysis legislation that would turn that card into a trump. But their position with respect to particulates is much less defensible than that of Greenpeace on Brent Spar, because the epidemiological evidence is so suggestive of significant risk to the population whereas mechanisms are poorly understood. The US legislature should reject the demands being made of science before regulatory change. □

Infected blood in Japan

Japan's regulatory system for drugs and other medical treatment needs to be completely overhauled.

THE scandal about the infection of thousands of Japanese haemophiliacs with HIV in the mid-1980s through the use of contaminated blood coagulants is coming to a head. After some unusually frank disclosures by the new head of the ministry of health, Naoto Kan, about the government's failure to take prompt action to prevent the spread of HIV through imported blood products that had not been heat-treated to kill viruses (see *Nature* 379, 663 & 760; 1996), attention is now focused on the individuals responsible. For the sake of the haemophiliacs infected and their relatives, it is essential that justice should be done and that all of those immediately responsible should be brought to task. But in this witch-hunt, Japan must not lose sight of the wider issue of the fundamental defects in Japan's system for drug approval that allowed the flawed decisions of these individuals to pass.

Japan has the highest per capita drug consumption in the world. Yet drug approval is regulated by only a handful of ministry officials, most of whom have no scientific qualifications. Beyond the ministry walls there also is a severe shortage of qualified experts to advise on newly emerging diseases such as AIDS, and, in Japan's hierarchical society, one or two senior academics can wield enormous power in the decision-making process.

There is no quick solution. Japan needs to upgrade its medical research organizations and produce more high-quality medical researchers. There is a crying need for an organization staffed by competent scientists to regulate drugs and new methods of medical treatment. Furthermore, Japan needs to ask why, despite having one of the highest rates of blood donation in the world, the country is still so dependent on imported blood products. □



Costs of satellite camera threaten to push Europe out of the picture

Munich. With less than four months to go before the contract for the Rosetta mission to the comet Wirtanen is due to be put out to tender by the European Space Agency (ESA), the question of who will control the closest and most spectacular images of a comet yet produced has still to be resolved.

A special meeting of representatives from ESA and its science programme committee (SPC) tomorrow (8 March) is expected to agree to a new 'announcement of opportunities' by ESA, inviting proposals to build a camera for Rosetta, the third of four 'cornerstone' missions being planned by ESA. But there is concern that, if non-ESA states provide substantial funding, the European agency may be deprived of much of the public credit for the mission.

Money is at the root of the problem. A sophisticated imaging system, able both to carry out a range of scientific analyses and to provide high-resolution photographs suitable for public relations purposes, was originally expected to be designed and paid for by a consortium of ESA member states, led by Germany. But these countries have now decided that they can no longer afford it.

Rosetta will meet up with the comet Wirtanen in deep space and accompany it for two years as the comet moves towards the Sun, to determine changes in both its nucleus and surface during the approach.

As with all ESA science missions, the costs of the launch and the spacecraft are paid for directly by the agency. But the instruments for the scientific payload, whose composition — apart from an imaging system — was agreed last month, are developed by outside scientists and paid for by national space and research agencies.

The proposal for an imaging system, called Osiris, capable of obtaining images in the visible, ultraviolet and near-infrared wavelengths, was given high scientific and technical ratings by Rosetta's independent peer review committee last autumn.

But the proposal was not given a top priority rating. Some claim this may have been because of hopes at the time — since dashed by the agency's financial difficulties — that ESA would provide a camera covering visible wavelengths, or perhaps because the navigational system, which controls the spacecraft's orbit around the comet through simple imaging, might be upgraded to take high-resolution colour photographs.

Furthermore, much of the infrared and ultraviolet range covered by the proposed

imaging system is already covered by other instruments on the payload. In these circumstances, Germany, which had already agreed to contribute significantly to other instruments with higher priority ratings, backed out of financing Osiris. So did the other three main partners, France, Italy and Belgium.

According to Marcello Coradini, secretary of ESA's Solar System working group, the SPC, which met last month to decide the

that the new pressure on national agencies may allow for "imaginative solutions" to be found, while the US National Aeronautics and Space Administration (NASA) may also have some interest in participating.

But this latter possibility has touched on sensibilities about who might take the credit for photographs resulting from the mission. Some ESA officials, as well as members of the SPC, are worried that the space agency could lose important public relations opportunities by allowing non-member states to have large stakes in the camera.

ESA

Bonnet believes that, whatever agreement is reached, "ESA must control the public relations aspects". Coradini agrees. "In years to come, the public will not remember the isotopic composition of the comet, but everyone will remember the images."

A second option, which might be adopted if an agreement on a new call for proposals cannot be reached tomorrow, would be a limited upgrade to the navigational system, which could be achieved without raising the total cost of the mission.

But Horst Uwe Keller from the Max Planck Institute of Aeronomy in Lindau, principal investigator for Osiris, says this would be a "disaster" for the mission, as it would restrict imaging to the visible range.

Gernot Hartmann, head of the extraterrestrial science department of DARA, the German space agency, which would have shared the German contribution to Osiris with the Max Planck Society, declines to comment on reports that Germany is studying ways it might provide support for a fully European imaging system.

But he admits that he would find it "awkward" if the response to a new announcement of opportunity for the imaging system camera "was not clearly a European effort".

Alison Abbott

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Eye on publicity: Europe might lose out on public relations if Rosetta's imaging system is paid for by other contributors.

final composition of Rosetta's payload, has put ESA in a "Catch-22" situation.

The committee confirmed that the cost of the mission should not exceed the ECU635 million (US\$800 million) in 1995 prices agreed for each cornerstone mission, ruling out the possibility of ESA providing a camera from its own budget. But it also insisted the scientific information from an appropriate imaging system should not be lost.

The programme committee has put forward two suggestions. Its preferred option is that a new 'announcement of opportunities' be put out as soon as possible for a scientific imaging system, built and operated by the space science community, and paid for by national agencies.

Although a previous announcement, only last year, failed to produce a viable proposal, Roger Bonnet, ESA's science director, says

Cuts put standards at risk, say UK universities

London. University vice-chancellors in Britain warned last week that standards in higher education could fall, following a 5 per cent reduction in their government funding for 1996/97. Some of the larger research universities have had to shoulder significant cuts to save smaller and newer institutions from unmanageable reductions in income.

Under the reductions, announced last week by the Higher Education Funding

Council for England (HEFCE), institutions such as the universities of Cambridge, Oxford and Imperial College, London, as well as some medical schools, will see their government funding cut by between 3 and 4.5 per cent. Funds for smaller universities and colleges, by contrast, will drop by between 1 and 2 per cent, while a few specialist institutions, such as the Wimbledon School of Art, will have their funds increased. □

US cancer institute shifts funds to external research

Washington. An advisory board to the US National Cancer Institute (NCI) has welcomed a decision by the institute to boost its funding for outside investigators by \$70 million in the current fiscal year, which began last October. Nearly half of this money is being taken from the institute's existing intramural activities — a move which, not surprisingly, is raising concern among NCI scientists whose projects may be affected.

The new plan is the work of Richard Klausner, the recently-appointed director of NCI, who presented it to members of the National Cancer Advisory Board (NCAB) last week. Describing insufficient research opportunities for outside investigators as "the problem within the core of our enterprise", Klausner promised that he would do "whatever it takes, as long as we actually have the money, to make opportunities available [to such researchers]".

The advisory board does not have the power to veto Klausner's plan. But it raised few questions and did not dissent from the strategy — reflecting the mood at the cancer institute, where the relatively high level of intramural, compared to extramural, spending has recently come under fire.

Last year, for example, a panel of prominent cancer researchers recommended sharp cuts in the level of research funds spent intramurally by the NCI (see *Nature* **375**, 267; 1995). The institute's 1996 budget of \$2.25 billion makes up almost one-fifth of total spending by the National Institutes of Health (NIH). Of this figure, 19 per cent goes to intramural research, compared to an average of 11 per cent at the NIH's 23 other institutes and centres.

Klausner's plan therefore represents "a tremendous advance", says Barbara Rimer, an oncologist at Duke University Medical School in North Carolina, and chair of the advisory board. "For the first time in years we have the opportunity to advance science in a dramatic way by enabling many more good ideas to get funded."

Klausner has already put in place the mainstay of this plan: an increase in

the coveted 'RO1' investigator-initiated research grants for basic and clinical research projects by outside scientists. According to Klausner, in the six months since he has become director, 23 per cent of applications for such grants have been funded, compared to only 15 per cent previously. This new level of funding will continue.

Furthermore, in a move that reflects recent calls for more support for clinical researchers, such researchers who fail to rank in the top 23 per cent, but fall within the next 10 per cent, will get an immediate second chance at being funded, provided that they respond promptly to assessments of their proposals. Non-clinical researchers will also be given the same second chance under the new process, called 'accelerated executive review', but only if they fall within the next 4 per cent of applicants.

Klausner calls the current process, in which the amendment and re-submission of applications can often take two years, "a particular problem for patient-oriented research". He told the board that the new, more flexible procedure sends "a very powerful message that we are interested in patient-oriented research". Klausner added that the NCI has set aside sufficient money to fund 40 to 50 per cent of those grant applications that qualify for the immediate second chance.

But the news of additional support for extramural projects has distressed some NCI scientists, who expect to see contract work cut by \$25–\$30 million to help pay for it. "All of us feel a bit like mushrooms — kept in the dark and fed you know what," says one laboratory chief at the auxiliary NCI campus at nearby Frederick, Maryland. According to Klausner, about \$10 million will be cut from contract research at the Frederick campus, whose current budget is about \$150 million.

"Obviously we're all anxious. We don't know where they're going to find the money," adds another scientist at the campus, the Frederick Cancer Research and Development Center (FCRDC), which provides support services for NCI AIDS and cancer research. The Center's senior manager, Joseph Mayo, says that the cuts will not require lay-offs, but will be absorbed through attrition among the approximately 350 NCI employees and 1,500 contract workers at FCRDC over the next two years.

Klausner claims that much of the new money has been found not by cutting programmes but by improving efficiency. NCI officials say that the remaining \$40–45 million of the \$70 million will come from the 5.7 per cent increase that Congress has granted NIH in 1996. **Meredith Wadman**

International biotech centre opens in Delhi with plea for cash

New Delhi. The final step in establishing the International Centre for Genetic Engineering and Biotechnology (ICGEB) took place in February with the opening of a new building for the centre's laboratories in New Delhi, after occupying temporary accommodation for the past nine years.

The centre's other base at Trieste in Italy moved into new laboratories last year. "With both components now installed in permanent premises, we have concluded the process of setting up the ICGEB," said Ambassador Sergio Cattani, representing the Italian government, at the formal inauguration.

But one major task is still left — ensuring that the centre will be able to attract the financial support that it needs to run its research programmes.

Until now, the two host countries, Italy and India, backed by donations from international agencies, have met the centre's expenses of US\$12 million a year. "Other member countries will have to start paying from 1999," says Arturo Falaschi, the director of ICGEB. "That is the deadline."

Handing over the 10,000-square-metre laboratory complex to ICGEB, Bhubanesh Chaturvedi, India's science minister, said his government had fulfilled its commitment to the international scientific community. While confirming India's continued support, he called upon other member countries to "urgently make their contributions to sustain the centre's excellence". The governments of 35 countries have so far ratified agreements with the ICGEB.

ICGEB, with its structure of twin laboratories, was set up in 1986 by the United Nations Industrial Development Organization (UNIDO) to ensure that developing countries are able to share the benefits of biotechnology. This objective has been achieved, says Falaschi, with about 230 scientists from 28 countries now working at the centre's two laboratories, and 400 more receiving short-term training each year.

Since February 1994, the centre has operated independently of UNIDO. "Now we are really on our own," says Falaschi, adding that "once all the member countries start making contributions, the centre will have no problem on the money front."

According to Falaschi, important contributions from the New Delhi laboratory include a new peptide vaccine for hepatitis B, and the isolation of a gene able to impart viral resistance to crops. An inexpensive AIDS diagnostic kit developed at the centre was put on the market in India last month, and will soon be produced locally in Nigeria, Sudan and Morocco.

K. S. Jayaraman

JET heads for reprieve

Munich. The European Parliament last week gave its backing to a European Commission proposal to extend the Joint European Torus (JET) nuclear fusion project for an extra three years. Final approval for the extension, which would allow JET to run until the end of 1999, is expected to be given by the European Council of Ministers in the next few weeks. The commission pays 80 per cent of JET's annual budget of around ECU80 million (US\$103 million). □

Congress cool on research funding proposals

Washington. A plan to centralize and streamline the allocation of US research and development (R&D) funding received a frosty reception in Congress last week, when members of both political parties on the Science Committee in the House of Representatives attacked its most important proposals.

The plan was put forward last November by a panel of the National Academy of Sciences chaired by Frank Press, former science adviser to US president Jimmy Carter (see *Nature* 378, 426; 1995). It would exclude \$30 billion of military development work from the science and technology budget, and require Congress to consider the remaining budget as a single entity.

Both George Brown (Democrat, California), the senior Democrat on the Science committee, and Vernon Ehlers (Republican, Maryland), a former physicist who wields growing influence on the Republican side, attacked the plan in a hearing on 28 February. They argued that the R&D budget would be harder to defend if it was shorn of the military development component, and that an additional step in the congressional budget process was unnecessary.

But Robert Walker (Republican, Pennsylvania), chairman of the Science Committee

until his retirement this November, welcomed the Press report. He said he would continue to push for the central consideration of the research budget by the Congress.

Walker nevertheless asked Press why the \$30 billion should be excluded from the \$75-billion annual R&D budget "when it uses so much of the nation's scientific resource and talent". Press replied that the military development work "is not R&D as most countries of the world define it".

But Ehlers said that while this proposal might be good logic, it was bad politics. "It's very logical and rational, but I have to worry about its political ramifications," he told Press, adding that a single research budget could be "a very tempting target" for budget-cutters.

Ehlers also pointed to some seldom-noted advantages of the existing appropriations process. There is, for example, "a lot of specific expertise in Congress about specific areas of the [research] budget". Ehlers acknowledged that he had been influenced by criticisms of Press's proposals made by D. Allan Bromley, dean of engineering at Yale University, who had been President George Bush's science adviser.

Meanwhile, Democrats on the committee focused on the implicit assumption of the

Press report that research budgets will decline. Brown said that the report failed to say whether its objectives could be met under current budget plans. "There comes a point when the whole research enterprise is at risk, and the problem cannot be solved



by better resource allocation," he said. Brown also expressed unhappiness at Press's handling of research allocation issues.

Asked directly if the objectives could be met if civilian R&D spending was cut by 33 per cent by 2002, as has been projected by the American Association for the Advancement of Science (AAAS), Press said "no", before adding: "I think the AAAS may be backing off from that projection, in the light of recent events." The AAAS denies that it is doing any such thing (see box, below left).

Connie Morella (Republican, Maryland), chair of the technology subcommittee, joined Democrats in expressing concern about Press's criticism of technology transfer programmes. The report criticized, in particular, the Advanced Technology Program, administered by the National Institute of Standards and Technology in Morella's district. Press said he had meant only that the programme should be thoroughly evaluated in comparison with other programmes.

Walker, the committee chairman, has warmly welcomed the Press report for two reasons. First, it echoes his own view that Congress should do more to consider the science budget across all agencies. Second, Press, rather than ignoring his congressional brief and simply asking for more money, chose to confront a new fiscal environment in which difficult choices must be made.

But Democrats think that Press went too far in tailoring his report — which was originally requested by senators of both parties — to the tastes of the new Republican majority. There is a chance, albeit a slim one, that the Democrats will regain control of the House next November. If they do, says Brown, the Press report will end up "in some obscure file somewhere".

Colin Macilwain

US politicians play the numbers game

Washington. Last July, the American Association for the Advancement of Science (AAAS) announced to a startled world that it expected Congress to cut non-defence research and development (R&D) by one-third by 2002. Since then, the number has assumed a life of its own. Most recently, it was used by Al Gore, the vice president, to attack Republican science policy (see *Nature* 379, 570; 1996).

So when Frank Press — wrongly, as it turned out — told the House of Representatives Science Committee last week that the AAAS was about to "back off from that projection", Republicans could scarcely conceal their glee. Robert Walker, chair of the committee, told the hearing he was "delighted", adding that he had been "disappointed that [the AAAS] used it in the way that they did".

After the hearing, Walker denied that his statements were an attack on the AAAS, which he said had "picked the figure up" from George Brown (Democrat, California), who made a similar analysis last spring. In fact, the AAAS got its figures from the seven-year 'concurrent budget resolution' passed by both houses of Congress last June. It found that the total non-defence R&D would

fall from \$34.2 billion in 1995 to \$28.5 billion in 2002. Making reasonable assumptions about inflation, this would be worth \$22.9 billion in 1995 dollars, a reduction of 32.9 per cent.

"We stood by the analysis then and we stand by it now," says Al Teich of AAAS. "But since then [last June] the Congress has been a lot kinder to non-defence R&D than it said it would be." For 1996, for example, appropriations committees have not made anything like the \$4-billion cut in R&D spending that the seven-year plan demanded.

The appropriations process for 1996 — due to be completed by 1 October 1995 but not yet finished — proved just how tough it was for even a freshly-elected, anti-spending Congress to cut actual programmes. This year, with elections pending, many observers say cuts will be even tougher to make, and spending could diverge still further from the harsh targets in last year's budget resolution.

This pattern could be repeated in future — or reversed. After all, both Congress and the Clinton administration are pledged to balance the budget by 2002. Perhaps the White House will now publish its own figures, explaining how this will be achieved.

C. M.

Britain may set up genetics advisory body...

London. The British government, in a significant shift from a position adopted only two months ago, has agreed to discuss setting up a transdepartmental advisory committee to monitor both the commercial applications and ethical implications of human genetics.

In doing so, the government appears to have agreed with members of the House of Commons Select Committee on Science and Technology that the existence of such a body might help to reassure the public that the potential dangers of the applications of genetics were being adequately considered.

The new move was announced last week by Stephen Dorrell, the Secretary of State for Health, in reply to questions about the government's response to a report on human genetics published by the select com-

mittee last July, after a seven-month inquiry.

Two months ago, in its formal response to the report, the government rejected its main recommendation (see *Nature* 379, 195; 1996). This was that a Human Genetics Commission be set up, with broad-ranging powers both to monitor and to regulate activities ranging from genetic screening to the way in which genetic information is used by insurance agencies.

At the time, the government claimed that the various aspects of the applications of human genetics were already satisfactorily covered by a range of advisory groups on topics such as gene therapy and genetic engineering experiments, and that a new over-arching commission would merely create additional bureaucracy.

But, in a rare parliamentary move, the all-

party committee was sufficiently frustrated by this response to call a further series of hearings designed explicitly to give various witnesses an opportunity to challenge the government's response.

Two weeks ago, for example, Sir David Weatherall, director of the University of Oxford's Institute of Molecular Medicine, and one of Britain's leading geneticists, told the committee that "in a very fast moving and complex field like genetics, some body is needed that is constantly monitoring things — if only because of their unpredictability — and is also able to reassure the public".

Last week, Dorrell told the committee that, having listened to the reactions to the government's response from both scientists and legislators, he had now accepted that the current monitoring arrangements might be insufficient to persuade the public that a broad perspective was being maintained on the commercial and ethical consequences of modern genetics.

As a result, he said, he had asked officials in his department to set up a review of aspects of the committee's proposal dealing with a body that would coordinate thinking across government of the direction and implications of developments in genetics.

Dorrell said that, given the specialist knowledge involved in many aspects of the

IMAGE
UNAVAILABLE
UNAVAILABLE
COPYRIGHT
REASONS
REASONS

of applications of genetics, he was reluctant to abolish the existing advisory committees — or to turn them into sub-committees of a centralized body. Nor did he accept the committee's suggestion that a new body should be given a regulatory role, suggesting that "both government and Parliament would want to decide themselves how to react to the advice on offer".

This itself has frustrated some committee members, who have been pushing for a commission with the power to lay down strict guidelines to prevent abuses of genetic information in activities such as health screening, insurance and employment.

But Dorrell, in contrast to the government's earlier statement, acknowledged his concern that "at present, there are a range of [advisory] activities going on across the board without any coherence". A key role of an "over-arching" advisory committee would be "to catalyse thought on the implications of [modern genetics] in the form of advice to the government" and, in doing so, to "lead public opinion" on the issues involved, he said.

David Dickson

...as ethics report clears xenografts

London. Britain's main bioethics advisory group has given the green light to xenotransplantation — the transplantation into humans of organs taken from other animals — and endorsed the use of pigs, rather than primates such as chimpanzees and baboons, as the most acceptable source of such organs.

At the same time, the bioethics group points out in a report published in London last week that there is still considerable uncertainty about the potential hazards of xenotransplantation.

No clinical trials of the technique should be permitted, it therefore suggests, until the government has set up a new Advisory Committee on Transplantation, and this committee has given its approval to the proposed trials.

The report has been produced by a ten-member working party set up by the Nuffield Council on Bioethics. It comes at a time of growing interest in xenotransplantation, not only because of the growing gap between the number of patients awaiting transplants and the supply of human organs, but also because of hopes that some of the main scientific hurdles — such as the 'hyperacute rejection' of pig organs by humans — may be close to resolution through genetic engineering.

Last year, for example, British scientists in the small company Imutran, based near Cambridge, announced that, having successfully bred pigs containing a human gene capable of switching off the production of the protein responsible for the hyperacute reaction, it confidently expected to begin clinical trials involving the transplantation of hearts from such pigs into humans before the end of 1996 (see *Nature* 377, 185; 1995).

The working group, which was chaired by

Albert Weale, professor of government at the University of Essex, concludes that the development of xenotransplantation "should continue", provided it is subject to "rigorous regulation" to ensure both the protection of potential human recipients and adequate care of the animals involved.

Given both ethical concerns about using primates such as chimpanzees which are widely seen as closer biologically to humans than other animals, and the potential threat of extinction that could result from the wide-scale use of primates for transplant organs, the working group suggests that "non-primate species should be regarded as the source animals of choice". Indeed, it says that the breeding of pigs to supply organs for transplantation "would be ethically justified".

But the working group also concludes that the risks associated with the possible transmission of infectious diseases, including, for example, the spread of currently unknown animal viruses in humans that might be difficult to control, "have not been adequately dealt with". In the light of such uncertainty, it recommends a cautious approach, in particular stating that at present "it would not be ethical to begin clinical trials of xenotransplantation involving human beings".

The group says that standards and mechanisms for monitoring xenograft recipients and for action to be taken in case of disease transmission "should be in place before human trials begin". It suggests that the proposed advisory committee should be responsible both for developing these procedures and ensuring that they are properly applied. But it also admits that the setting up of such a committee is likely to encounter some reluctance from the UK Department of Health.

David Dickson

No takers so far for NASA science institutes

Washington. A full year after announcing plans to transfer some of its research operations to privatized science institutes, the US National Aeronautics and Space Administration (NASA) has yet to find anyone willing to sponsor an astrobiology institute associated with the agency's Ames Research Center in San Francisco.

Furthermore, although the team responsible for setting up the institutes issued its final report last week, critics charge that the overall concept still lacks detail. "We still don't have enough specific information to make any informed judgement," says Mary Jane Osborn of the University of Connecticut Health Center, a member of the National Research Council's Space Studies Board, which was briefed on the plan last week.

The science institutes were initially conceived to save money. But they are now seen as a way for the space agency to improve the quality of its science and strengthen its ties with the outside research community (see *Nature* 378, 121; 1995).

The NASA study team headed by Alphonso Diaz, now deputy director of the Goddard Space Flight Center, recommended last November that three of the 11 proposed institutes were ready for implementation: a biomedical research institute affiliated with the Johnson Space Center in Houston; the Ames astrobiology institute; and a microgravity institute at the Lewis Research Center in Cleveland, Ohio.

Of the three, planning for the biomedical institute has progressed the furthest. Diaz says he hopes to hold a 'bidders' conference' this spring for potential sponsors in the university and non-profit sector. But many questions remain, including whether institutes must be located in the same area as the NASA centre, and what the chain of command would be between NASA headquarters, the institute, and Johnson. At present the relationship is fuzzy, says Osborn.

Particularly worrying is the question of peer review. Life science at NASA, plagued in the past by charges of shoddy research, has only recently begun to shed that image thanks to rigorous peer review procedures instituted at the agency's headquarters. To shift such responsibility back to the centres or to an institute would be "a very unacceptable outcome", says Osborn.

At Ames, hundreds of scientists currently employed by the federal government would probably be transferred to the new institute. NASA has proposed that new legislation should protect these civil servants' benefits, preventing them from being barred from working for the institutes by laws restricting post-government employment.

To do this, the agency has proposed language to be included in its 1997 authorization bill, and Robert Walker (Republican, Pennsylvania), chairman of the House

Science Committee, has also included a provision easing employment restrictions in the draft of a space commercialization bill he plans to introduce later this year.

But personnel issues are not the only reason that no-one is beating down NASA's door to take on the astrobiology institute at Ames. The work of the institute would cut across three disciplines: space science, life science and Earth science. Some potential sponsors have expressed interest in pieces of each, says Diaz; but "very seldom did we find that people were interested in the full breadth of what we're offering".

NASA's answer is to hold a workshop in

the summer to show off the range of science operations at Ames. While Diaz still hopes to sell the whole facility to a single buyer, NASA will consider selling individual lots. "We don't want to put out a requirement that nobody responds to," he says. "But we obviously don't want to break it up into so many small pieces that it creates a bureaucracy to integrate it all."

In short, the agency is trying to execute something that has never been tried before. "I don't think it's a matter of NASA trying to evade the questions," says one Congressional staff member. "I really don't think they know at this point."

Tony Reichhardt

Untethered satellite still produced data

Munich. Silvano Casini, chief administrator of the Italian Space Agency (ASI), last week defended the agency's ill-fated tethered satellite mission. Despite the dramatic loss of the satellite after its tether snapped as it was being unreeled from the space shuttle, Casini said that many of the scientific aims of the mission had been achieved.

But the bad publicity remains a further unwelcome burden for ASI — as well as its partner in the mission, the US National Aeronautics and Space Administration (NASA) — just as the Italian agency is struggling to throw off an image of inefficiency and wastefulness.

The main aim of the mission was to investigate the creation of an electrical current along a 20-km conductive tether, attached at one end to the space shuttle, as it was dragged through magnetic fields by a satellite moving in a slightly higher orbit, with the electrical circuit being closed by electron guns in the shuttle's payload bay discharging the electricity back into the ionosphere.

At least the test of this idea was successful. Nearly the whole length of the tether had already been unwound, at 2.2 metres per second, before it snapped within the deployment boom, giving enough time for scientists to observe the predicted flow of electricity. A voltage of 3,500 volts and a current of 480 mA were measured shortly before the break.

A second aim, to test the dynamic behaviour in orbit of a large rigid system 20 km in length, failed. But there was some success with other experiments conducted by the satellite not directly related to the tether, exploiting instead the satellite's position in a little-explored orbit which is unstable for free-flying satellites. As radio contact with the satellite was restored after the tether had broken, and was maintained for a

few hours a day for four days before the satellite's batteries died, some data describing the plasma environment in the upper ionosphere magnetic field were successfully generated.

Yet despite the relative scientific success of the mission, and uncertainty about how — and if — blame should be allocated, the perceived failure of the experiment in the eyes of the Italian public is a further cross for ASI to bear at a time when it is under considerable political pressure. Indeed, some politicians immediately questioned the value of Italy's space programme, given the country's major economic difficulties.

ASI, long criticized for its inefficiency, is in the hands of a temporary administration, whose mandate to restructure the agency expires in June. A new law governing its structure and activities was expected to be in place by then. But the recent call for elections in late April means that this will be virtually impossible to achieve.

Meanwhile, ASI may have to face yet another problem. Its X-ray satellite, SAX, which has been subject to continual and controversial delays, is being prepared for launch on 29 April. But a group of specialists from the European Space Agency, which has been contracted to review the readiness of the ground segment of the mission, is expected to report this week that this part of the mission could be significantly improved.

ASI has to decide whether to go ahead with the launch without the improvements, or to delay the mission until the next possible launch date in October 1997, eight years later than originally planned. Italian scientists have continually criticized the project for its escalating costs and delays, which they claim, have made its science outdated.

Alison Abbott

Plans for N-test simulators trigger reaction

Paris. A shift in the focus of France's nuclear weapons research programme from developing nuclear weapons to simulating nuclear explosions is prompting a major reorganization of the Direction des Applications Militaires (DAM), the defence arm of the French Atomic Energy Commission (CEA), which announced last week that it is to shed one-fifth of its 5,700 staff over the next decade.

But France's decision, like that of other nuclear weapons states, to shift its weapons research to simulation techniques remains controversial. In particular, critics such as the Nuclear Control Center in Washington DC claim that such techniques — and lasers in particular — could be used to design new nuclear warheads.

The shift in research goals follows President Jacques Chirac's recent decision both to cease nuclear testing and to make major cuts to France's nuclear weapons programme. Addressing the nation on a major reform of the armed forces, Chirac confirmed that France would retain its nuclear deterrent as an ultimate defence. But he added that the scope of its nuclear weapons activity would be reduced to a minimum.

The brunt of the cuts will be borne by DAM, which will close two research centres, at Limeil and Vaujours, and regroup its activities in its four remaining laboratories, at Valduc, Ripault, Bruyères-le-Chatel and Le Barp. According to Jacques Bouchard, director of DAM, however, most of the cuts will fall on activities connected with developing nuclear weapons, and not on research.

Indeed, while overall staff numbers are being reduced, Bouchard says that he will still be able to increase the recruitment of scientists to nuclear simulation research, as many DAM staff will retire soon.

Similarly, France's investment in simula-

tion is set to grow. It will spend FF150 million (US\$30 million) this year on supercomputers, for example, and will invest a further FF750 million on Airix, a powerful X-ray machine used to study the behaviour of heavy metals during implosion of chemical explosives.

The Centre d'Etudes Scientifiques et Techniques d'Aquitaine (CESTA) at Le Barp, near Bordeaux, will be the lynchpin of DAM's new strategy. It will house the FF6-billion 1.8-megajoule laser being built as part of French plans to carry out nuclear weapons research without having to carry out nuclear tests of its hydrogen bombs (see *Nature* 375, 6; 1995). To build the *Megajoule* laser, DAM is collaborating with the Lawrence Livermore National Laboratory in California, where a similar facility, the National Ignition Facility (NIF), is to be built.

Bouchard dismisses claims that the simulation techniques might be used to design new warheads. He argues that simulation will be used only to renew existing weapons in the stockpile and to allow minor modifications to warheads, for example, to fit improved missiles. Simulation is needed, he says, to maintain a critical mass of competence in nuclear weapons that will in turn ensure that the reliability of the existing stockpile is maintained.

Indeed, many US weapons experts, including Richard Garwin, IBM Fellow Emeritus at the Thomas J. Watson Research Center in Yorktown Heights, New York, agree that while rudimentary weapons could be designed without testing, nuclear weapons states have no interest in adding such weapons to their stockpile, as they already possess reliable high-performance weapons.

Another concern raised by Theodore

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Blast from the past? France plans to close nuclear testing sites at Fangataufa (above) and Mururoa.

Taylor, a former nuclear weapons designer at the Los Alamos National Laboratory, is that the development of lasers capable of triggering thermonuclear fusion itself poses a new threat of nuclear proliferation.

Taylor claims, for example, that such lasers could be an important step towards designing pure fusion weapons, or directed-energy weapons such as high-power microwave beams, and has called for a ban on the development of the technology.

Bouchard argues that the *Megajoule* laser can in no way be considered a weapon. "It is simply a big physics instrument," he says. But other critics, such as Venance Journé, from the Centre International de Recherche sur l'Environnement et le Développement (CIRED) in Paris, still complain about what they claim is a lack of adequate information about the nature of planned experiments using the laser.

The secrecy surrounding such projects, she adds, highlights a wider problem, namely the failure of the French government to allow independent scientific experts sufficient access to classified information. "Those who know, don't talk, and those that talk, don't know" is how one observer paraphrases what he claims to be a lack of informed debate in France about the necessity of resuming nuclear testing.

Bouchard says he agrees that sufficient information about nuclear weapons should be made available to allow "real discussion between specialists". At the same time, however, he claims that it is dangerous to provide "too much information in such a sensitive area", arguing that this itself could increase the risk of proliferation.

But Journé contests this, pointing out that the United States allows independent experts access to classified data without jeopardizing national security. For example, a panel of experts convened under the JASON programme last year studied and rejected the case for small nuclear explosions, later releasing an unclassified executive summary of the full classified report (see *Nature* 376, 540; 1995). **Declan Butler**

France seeks to clean up nuclear image

Paris. President Jacques Chirac's announcement of a unilateral reduction in France's nuclear weapons capacity, issued on the eve of the first major summit between European and Asian nations, was clearly intended to restore France's international reputation following the damage done last year by its decision to proceed with a series of nuclear tests in the South Pacific. Ironically, Chirac is now presenting France as the champion of nuclear disarmament.

The reduction will include the dismantling of France's 18 one-megaton silo-based missiles on the Plateau d'Albion, and 40 mobile 80-kiloton Hadès missiles. Its nuclear deterrent will now consist primarily of a modernized fleet of four nuclear submarines, with a few

dozen aircraft. France will also stop producing weapons-grade fissile material, and will close its plants at Pierrelatte and Marcoule; it has stocks to meet anticipated needs for at least 50 years.

Chirac reiterated his commitment to supporting a ban on all nuclear tests, no matter how small (see *Nature* 376, 540; 1995). He also announced that France would close for good its nuclear testing sites at Mururoa and Fangataufa in the South Pacific.

In taking this action, Chirac has preempted other nuclear weapons states, which have also promised to sign the Comprehensive Test Ban Treaty (see *Nature* 376, 283; 1995), but which have not yet decided definitively to close their testing sites. **D. B.**

Scientist in Japan's HIV row resigns as university vice president

Tokyo. Takeshi Abe, one of the scientists at the centre of Japan's contaminated blood products affair, has resigned as vice president of Teikyo University. But he continues to insist he is innocent of wrongdoing in the events surrounding delays to the distribution of heat-treated blood products in the 1980s, and says he is resigning in order to avoid causing trouble for his university.

In 1983, Abe was head of a Ministry of Health and Welfare AIDS study group which endorsed the use by haemophiliacs of blood products that had not been heat-treated. Last month, a mother whose son died in 1991 after transfusions with HIV-contaminated blood filed an accusation of murder against Abe. She claims that 23 out of 48 haemophiliacs in his care had been infected with HIV after receiving similar products. □

Harvard and Biogen file patent suit

Boston. Harvard University has filed a joint suit in a Stockholm court with the biotechnology company Biogen Inc., both based in Cambridge, Massachusetts, against the Swedish-American pharmaceutical company Pharmacia & Upjohn Inc., for allegedly infringing the patent on a technique for creating human hormones.

Biogen is claiming \$141 million in damages from past sales of Genotropin, the growth hormone produced by Pharmacia & Upjohn, as well as 7 per cent of revenue from future sales of the drug in Sweden. Harvard has refused to make any public comment on the case, and a spokeswoman for Biogen declined to say any more than that the company intends to pursue the suit "vigorously".

The controversy dates back to research conducted in the late 1970s by a group headed by the Nobel prizewinner Walter Gilbert, of the department of molecular and cellular biology at Harvard. In 1978, the team developed a method for secreting the hormone insulin from inside bacterial cells. "This was the first hormone made by taking a natural gene and making it function in bacteria," says Gilbert. A statement released by Pharmacia & Upjohn confirmed that the company rejects the charge that it had breached any patent, and is not unduly concerned about the lawsuit. □

Team wins project for X-ray mission

London. A consortium of European organizations headed by the University of Leicester in the United Kingdom has been picked by the European Space Agency (ESA) to provide details to the scientific community of observations from a new X-ray telescope that will be launched in late 1999.

The eight-member consortium, drawn from institutes in France and Germany as well as the United Kingdom, will run the mission Survey Science Centre (SSC) for the £300-million (US\$460-million) X-ray Multiple Mirror satellite mission (XMM), which has been dubbed the 'Hubble Telescope' of X-ray astronomy.

The SSC team will be involved in developing the science analysis software for XMM and in processing the observations once the data begin to arrive. □

A male majority in phase-3 trials

Washington. The first data from a new demographic tracking system at the National Institutes of Health (NIH) show that about 35 per cent more males than females are used as subjects in phase-3 clinical trials. The full significance of these data will require future studies to provide comparative numbers. But according to some observers, the disparity is less than had been expected, given that the NIH only recently required that women and minorities be represented in the clinical trials which it funds.

The preliminary data, released by the NIH's Office of Research on Women's Health (ORWH), cover 400 mixed-gender phase-3 trials funded by the NIH in 1994. Of 159,904 trial subjects, males represented 57 per cent (90,483) and females 42 per cent (67,149). A further 2,272 subjects, making up 1.4 percent of the total, were not identified by sex.

"When you see data like that, the question is: are there enough women?" says Eugene Hayunga of the ORWH. The answer, he adds, depends on the science in question. At least one women's group has welcomed the results. "Considering where we started from, this is great progress," says Phyllis Greenberger, the executive director of the Society for the Advancement of Women's Health Research, adding that the mere fact that the data are now being kept represents significant progress over past procedures. □

South Pacific nuclear study begins

London. An 18-month study into the radiological situation and long-term impact of nuclear testing on the South Pacific atolls of Mururoa and Fangataufa begins this week. It will be coordinated by the International Atomic Energy Agency (IAEA), and is being carried out at the request of the French authorities.

The study will be carried out by an international advisory committee of experts drawn from 10 countries, as well as officials representing organizations such as the South Pacific Forum, the World Health Organization and the European Commission. The French authorities have agreed to provide the IAEA with relevant information and data. Analysis of terrestrial and marine samples will be carried out at the IAEA laboratories in Seibersdorf, Austria, and in Monaco and elsewhere. □

Nuclear cargoes 'need stricter rules'

London. Environmental groups are proposing stronger rights for coastal states to control the passage of ships transporting radioactive cargo, as well as the strengthening of safety measures for ships carrying radioactive materials.

Edwin Lyman, scientific director of the Nuclear Control Institute in Washington DC told a meeting of the International Maritime Organization in London this week of the need for "more stringent safety measures" to compensate for weaknesses in the packaging of radioactive waste, plutonium and spent fuel being carried on ships.

Jon Van Dyke of the University of Hawaii Law School told the meeting that a "more restrictive regime" was needed to prevent ships carrying "ultrahazardous" cargoes from setting sail without notifying and consulting coastal states on their routes. □

Leonardo's science studies on view

London. A selection of 100 drawings by Leonardo da Vinci, the fifteenth century architect to European aristocracy and military engineer, went on display at the Queen's Gallery at Buckingham Palace in London, last week. The drawings are part of the private collection of Queen Elizabeth II.

In addition to designing the first parachute and studying the movement of birds and fish in order to design aeroplanes and submarines, da Vinci studied the structure of the muscles and bones of the human body — dissecting some 30 corpses in the process. He also studied the heart and speculated about the circulatory system one century before the English physician, William Harvey. However, da Vinci kept most of his ideas and drawings to himself. The drawing above, like most of da Vinci's studies, is annotated by notes in his trademark cryptic mirror-writing. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Risk: a suitable case for analysis?

Is scientific risk assessment the best way to make decisions about environmental threats? Republicans in the US Congress, and many in industry, think it is. Others — including environmentalists — are less convinced.

LAST month, the Environmental Protection Agency (EPA), won an undesired accolade. The US government agency, which is charged with protecting the public from environmental hazards, won the "Panic at the Drop of a Rat" award from the Competitiveness Enterprise Institute (CEI), for its reaction to a scare over dioxins.

In 1983, the EPA responded to an accidental spillage of dioxin in the town of Times Beach, Missouri, by evacuating and demolishing what had been a community of 2,200 people. The CEI — a stridently anti-regulation lobby group — says that its award "symbolizes the all too frequent phenomenon of government policy and action taken on the basis of shoddy or insufficient scientific evidence".

The fate of Times Beach has become a useful symbol for business-funded groups such as the CEI in their battle to discredit the regulatory process of the US government. In this regard, it is a mirror image of the Brent Spar, the disused oil storage buoy whose proposed disposal in the Atlantic Ocean became a symbol for European environmental groups, who used its potency to help block the dumping last summer (see page 13).

The noisy clashes over such symbols have come to dominate many public debates on the environment. In cases such as Brent Spar, they have dictated environmental decisions too. So, over the past ten years or so, some policy-makers have sought a better way. To make decisions that will not become the subject of ridicule, like Times Beach, or collapse under public opposition, like Brent Spar, they have turned to risk assessment.

Risk assessment — or scientific risk assessment, as it is more grandly known — is an attempt by scientists and engineers to assess objectively a public health or other environmental risk. The assessment is then used by the decision-makers in the agency — the risk managers — in reaching an essentially political decision about how the risk should be dealt with.

Since its genesis in US attempts to quantify the safety of nuclear power plants twenty years ago, risk assessment has gradually evolved towards something approaching a

formal method with four stages:

- source/release assessment: to estimate the likelihood and scale of a potential release of a hazardous material;
- exposure assessment: to estimate what will reach whom;
- dose-response assessment: to estimate the impact of exposure on human health;
- risk characterization: to collate the above in a useful way.

All of these stages customarily involve great uncertainty, as the two examples on

The Brent Spar oil storage buoy (above), and the dioxin-exposed town of Times Beach (left), both subjects of hasty decisions which might have benefited from a deeper analysis of the complex choices facing policy makers.

the following pages amply demonstrate.

In the case of the Brent Spar oil storage buoy, a form of risk assessment has been initiated by the British government after public opinion — in the shape of a vigorous campaign by Greenpeace and a Europe-wide consumer boycott of the Shell Oil Company — had blown the company's previous position out of the water.

When the EPA considers new regulations of small, airborne particles (see next page) it has been required to conduct a thorough risk assessment, despite an extreme dearth of scientific information.

Proponents believe that the process can still be useful because it is the only objective technique for measuring and comparing risks. Also, they say, it can help to focus scientific research, as well as environmental

regulation, on areas of greatest need.

The practitioners of risk assessment are trained scientists and engineers, many of whom were drawn from research careers by a demand to quantify risk. In the United States, combative relations between corporations and government have often nurtured that demand, and risk assessment blossomed early in the United States. Today, the US-based Society of Risk Analysis has some 2,000 members, including around 200 in Europe and 200 in Japan.

According to Michael Firestone, a science adviser in the EPA's office for pesticides and toxic substances, the risk assessor's job is "to provide as much information as possible" to the risk managers, their seniors in the agency who will make actual regulatory decisions. People attack the EPA's risk assessment procedure, says Firestone, when what they are really opposed to is its risk management decisions.

Nonetheless, risk assessment has emerged as a kind of surrogate battlefield in the war between the industrial and environmental lobbies over regulation. The environmentalists were suspicious of scientific risk assessment in its early days, believing that it would be excessively influenced by industrial interests, and used to block sensible, precautionary regulations, whose benefits could not be clearly quantified.

According to Debbie Stine, who has directed major studies of risk assessment at the National Academy of Sciences, "the environmental groups are much more accepting of risk assessment now than they were in the past". Today, the EPA makes regular use of risk assessment, and the focus has moved on to exactly how it is used.

With that in mind, Congress has made repeated attempts to legislate the practice of scientific risk assessment. Senator Bennett Johnston (Democrat, Louisiana) tried to pass a law to tighten up risk assessment practice and streamline it between agencies in 1994. But that proposal did not go far enough for the Republicans, who have played an increasingly powerful role since winning control of both houses of Congress

E.P.L./David Sims

On other pages

■ Questioning on particulates	11
■ Can risks be quantified?	12
■ Brent Spar: don't blame science	13
■ Hazards under the microscope	13
■ Who's now saying what	14

in November that year.

The EPA and other government agencies are required by current law to conduct risk assessment, but they are under no obligation to manage risks on the basis of its outcome. The EPA's duty to protect public health takes priority over its obligation to assess risks. The Republicans sought to change that, and make it extremely difficult for the agency to take regulatory action unless it could prove, through risk assessment, that such action was needed. A law to achieve this became the top priority of the science committee in the House of Representatives.

The law was passed by the House but was vigorously opposed by President Bill Clinton's administration and is now stuck in the Senate, where its sponsor, Bob Dole (Republican, Kansas) cannot quite muster the 60 per cent majority needed to proceed. The House law includes provisions to open up the risk assessment process, to subject that process to judicial review in some circumstances, and to require assessors to make and publish a "centralized" estimates of the risk they are assessing.

Opponents of the legislation say that none of these provisions should be taken at face value. Openness is a ploy to open the process still further to the influence of industrial scientists who, opponents say, will

be the only ones with the time and money to take full advantage of it. And judicial review will leave the regulator's work constantly bogged down in the courts, they say.

The requirement to make and publish 'centralized' risk estimates angers many risk assessors. At present, the EPA relies on risk estimates that are conservative from the point of view of public health. Where there is uncertainty, it veers on the side of caution. The bill would require it to take best and worst cases, and estimates in between, and come up with 'centralized' risk estimate.

"It's a meaningless term," says Terry Davies, director of the Centre for Risk Management in Washington DC. "The EPA is in the business of protecting public health." As such, says Davies, it needs to consider the worst cases, rather than the average risk.

Congressman Robert Walker (Republican, Pennsylvania), chair of the House Science committee and chief sponsor of the risk assessment legislation, denies that its objective is "paralysis by analysis". He concedes that there may be some supporters of his bill who just want to block regulation altogether, but says that his goal is to have "regulation based on a common sense approach" using "the best science".

Don Ritter, a former Republican congressman who now heads the moderately

conservative National Environmental Policy Institute, believes that some good has come from the argument over the legislation, as the issues were aired more thoroughly in Congress and the media, than ever before. "People think that because the risk assessment bills didn't pass, risk assessment itself took a beating," he says. "But risk assessment went from the arcane to the acceptable in the course of the debate. And the concepts of risk assessment were not rejected. They will be back, in one form or another."

Jim Wilson, a past president of the Society for Risk Analysis now with the Centre for Risk Management, says that scientific risk assessment is continuing to gain impetus regardless of the intervention of lawmakers.

Ultimately, in democratic states, decisions about what to do about disused oil platforms and particles in the air we breathe will be decided by politics, not science. Risk assessors are working diligently to develop techniques that will direct the best available scientific information into the political process. But their craft is in danger of being hijacked by those who are chiefly interested in ensuring regulatory outcomes to their liking. In those circumstances, it will be hard for risk assessment to win a central role in the way in which governments manage risk.

Colin Macilwain

Regulators face questioning on particulate rules

Washington. Risk analysis is increasingly accepted by US federal agencies as a way of injecting greater rationality into decisions on the regulation of environmental threats. At the same time, it has another function: as a defence for agency officials against those who claim that they are not adequately meeting their responsibilities.

Twenty years ago, for example, faced with the complaint that tiny particles of dust are a public health hazard (see panel over), the US Environmental Protection Agency (EPA) might have proposed, without much hesitation, tighter rules regulating their emission. But in today's more conservative America, with congress watching environmental regulation like a hawk, the EPA is proceeding with caution.

In the new political climate, it has been forced, by legal action, to review its air-quality standard for airborne particulate matter (PM). The EPA is now trying to construct a comprehensive risk assessment plan, in part to ensure that it has a defence against legal attack if it tries to make that standard more restrictive.

But the still-emerging risk assessment plan for PM — which itself will be reviewed by the agency's Clean Air Scientific Advisory Committee (CASAC) — is dogged by significant scientific uncertainty about which particulates damage human health and how they do it. Epidemiologists claim to be able to link particulates to increased sickness and death at exposure levels well below today's

legal threshold. But the exact nature of the risk remains to be established, and it may be years before science can settle the matter with any confidence.

Despite this uncertainty, the law demands action now. Particulate regulations come under the Clean Air Act, which requires the EPA to review its air-quality standards every five years. The last time the

agency did so, in 1987, it switched from a standard based on Total Suspended Particulates (TSP) to one based on PM₁₀ (particles smaller than 10 microns), arguing that particles larger than this do not penetrate deeply into the lungs.

But EPA missed its 1992 deadline for revising the standard. That prompted a lawsuit by the American Lung Association, as well as a court order compelling the agency to decide by November this year whether to stay with the current PM₁₀ standard, or to propose a new one, to take effect in June 1997. The court order is now driving the entire debate on particulates — at much too fast a pace, according to most experts.

It is CASAC's job to review the voluminous documents that EPA uses to build its scientific case for changing any standard. In meetings in August and December last year, the advisory committee criticized key sections in those documents in the particulates debate, and of the short timescale it had been given to review them. Indeed, CASAC members refused to endorse a preliminary recommendation by EPA staff that the agency should set a new standard for particles under 2.5 microns in size (PM_{2.5}).

George Wolff, the chairman of CASAC, and a scientist with General Motors Corporation's environmental and energy staff, told Carol Browner, the administrator of EPA, in a blunt letter that "the agency does not adequately justify the selection of PM_{2.5} as the best surrogate measure for the ►



Hazards ahead: but industry still insists that air emission rules are based on 'facts'.

Quantifying risk can be easier said than done

Do fine particles in the air constitute a health hazard? Many of the epidemiological studies carried out in the past several years suggest that they do, even at levels well below current US government safety standards. But scientific uncertainty, particularly about the physiological mechanisms involved, makes precise risk analysis difficult to achieve.

Data from dozens of American cities, as well as from eastern Germany, Beijing, Athens, São Paulo and other places, have consistently linked airborne particulate matter (PM) — which ranges from windblown dust to black diesel smoke — to respiratory problems and mortality due to cardiopulmonary disease.

If true, one explanation may be what is called 'harvesting', the fact that people already close to death are pushed over the edge during episodes of severe pollution. That in itself would be cause for concern. But a study published last year by Arden Pope of Brigham Young University, Provo, Utah, and colleagues was perhaps more disturbing.

Pope's team compared ambient particulate levels in 151 US metropolitan areas with mortality data for more than 500,000 adults enrolled in a prospective study for the American Cancer Society (ACS). After controlling for smoking, education and other factors, the authors concluded that death rates were as much as 17 per cent higher in cities with the relatively high concentrations of $PM_{2.5}$ (particles less than 2.5 microns in size).

Such results have captured the interest of health researchers, and also raised alarm among public health officials. But the particulates 'problem' still has a maddening number of scientific uncertainties,

starting with the question of which pollutant should be considered the culprit.

Particulates come in all forms and sizes — liquid and solid, coarse and fine. Epidemiological studies have tended to



What's your poison? Particulate threats to health can come from a wide variety of possible sources.

measure different things, sometimes even using surrogates such as sulphate aerosols when there are no data available for particulates in general. But that leaves uncertainty about whether it is the chemistry, or the size, of the particle that causes the problem.

The impact of particulates on health also tends to vary according to whether or not other air pollutants, such as sulphur dioxide, are present, making it difficult to untangle which factor is to blame. Even weather can confuse the picture.

Nevertheless, an independent review of six key studies, conducted last year for the Health Effects Institute, an organization based in Boston, Massachusetts, that is jointly funded by industry and EPA, generally confirmed the link between high particulate levels and increased death rates, even when different statistical techniques

and assumptions about weather and other pollutants were used.

But the authors, led by Jonathan Samet of the Johns Hopkins University School of Public Health in Baltimore, Maryland, nevertheless warned that, on the basis of a thorough re-working of data from Philadelphia, one of the cities studied, it was impossible to attribute observed effects to particulates alone as opposed to air pollution in general — adding yet further complications to any risk analysis.

To add to the difficulties, even the biological mechanisms behind particulates-related health effects remain largely speculative. Some researchers have focused on the acidity of particles, as this is known to affect pulmonary function. Others point out that the very smallest particles can penetrate deeply into the lung.

Anthony Seaton and his colleagues at the University of Aberdeen in Scotland, for example, proposed last year that such 'ultra-fine' particles inflame passages in the lung and stimulate the production of blood coagulants, causing fatalities among some chronic heart patients.

EPA researchers have also looked at the possibility that organic matter and metals, such as iron, attached to the surface of small particles might generate reaction products, which would interfere with the normal functioning of cell membranes and receptors in the lung.

But so far, these are only isolated hypotheses. The one aspect that all agree on is that more research is needed if toxicologists are to help solve a mystery that has been studied mainly by epidemiologists — and provide the data on which a useful risk analysis can be based.

T. R.

►causative agent [of health problems].”

Going even further, the full committee claims that the EPA documents had not provided “an adequately articulated scientific basis for making regulatory decisions concerning a PM”.

Faced with the embarrassing prospect of putting forward a new standard for particulates without the blessing of its own scientific review board, EPA has sought — and won — a series of extensions from the court to refine its arguments. The American Lung Association, which has been pushing for a decision, reluctantly agreed to the extensions, as a result of which CASAC will now be able to review EPA's PM documents this spring.

The advisory board wants to know in particular how the environmental agency will quantify the risks of particulates. In January, it asked EPA to decide how to carry out such

a risk analysis, and to indicate how death and disease rates would be reduced if a $PM_{2.5}$ standard were put into effect.

The agency's response, presented to the advisory board at a meeting in North Carolina last week, would commit the EPA to providing an estimate of risk for selected sample cities, rather than the entire country. The risk model would predict the health effects of different levels of particulates, based on 'concentration-response' functions derived from epidemiological studies.

But many uncertainties still remain. These include whether ambient particulates measured at monitoring stations correlate with the doses to which individuals are exposed, the role of other pollutants, and whether the use of base-line information on health effects that is not specific to the city in question skews the result.

According to EPA's plan, the risk analysis

would handle this built-in uncertainty by clearly stating and explaining the limitations and assumptions contained in the risk quantification. A 'confidence interval' would be given with all predictions of health effects for various levels of particulate.

Nevertheless, says EPA's draft plan, “due to the many sources of uncertainty inherent in such analyses, any risk results should not be interpreted as precise measures of risk”.

That answer may not satisfy CASAC, let alone politicians. Indeed, even if it was given more time, some members of the advisory board doubt privately whether EPA will be able to build an air tight case for a new standard when so many fundamental scientific questions remain answered. As a result, particulate regulations are likely to be left unchanged next January, and scientists — as well as risk assessors — sent back to the drawing board.

Tony Reichhardt

Brent Spar: when science is not to blame

Munich. In some environmental situations, scientific uncertainty is the biggest problem faced by any attempt to conduct a rigorous risk analysis. In others, the scientific arguments for or against a particular line of action may be relatively clear-cut; in such instances, the task faced by decision-makers of persuading a sceptical or hostile public of the legitimacy of their conclusions can be equally difficult.

Such was the situation last summer, when the British/Dutch oil company Shell was forced by an astonishingly effective Greenpeace campaign to abandon its plans to dispose of the Brent Spar oil storage buoy in the Atlantic Ocean. It seemed at the time that Goliath had been justly felled by David. Soon afterwards, as it emerged that Greenpeace had mistakenly used false information about the buoy's contents to swing public opinion, it was less obvious which side held the moral high ground.

Risk analysis may help to provide an answer. But, even after the publication — expected shortly — of the report of a scientific advisory body set up last year by the British government (see box) on the likely impact of different forms of deep-sea disposal of the Brent Spar, such analysis is unlikely to resolve the conflict over its fate.

Shell believed — and still believes — that the environmental risk posed by dumping Brent Spar in the ocean is inconsequentially small. The oil company had even commissioned some research which supported that view. But the public agreed with Greenpeace that the risk was conspicuously large. Could an extensive, independent and thoroughly scientific risk assessment process have changed their minds?

The saga started last February, when the British government approved Shell's request to dispose of the Brent Spar 2,400 metres beneath the Atlantic, at a site on the UK continental shelf called the North Feni Ridge. Shell says it had identified the site as its Best Practicable Environmental Option (BPEO), after balancing scientific and environmental considerations with safety, health and economic criteria.

Initially, according to Shell, it considered thirteen different options for abandoning or re-using the Brent Spar. The oil company chose deep-sea disposal on the basis of three main criteria — safety, cost and environmental impact. The alternative seriously considered was horizontal dismantlement on land: the buoy would be turned on its side at sea, its two damaged storage tanks repaired, and after a rough clean-out it would be towed to shore for dismantling.

But, according to a risk analysis carried out by Aberdeen University Research and Industrial Services before the disposal option was chosen, dismantling was four times as expensive as deep-sea disposal. Furthermore, the risk of a fatal accident was six

times higher, because of the labour-intensity and complexity of the process.

Shell's assessment of the environmental risk — which it thought would be negligibly small in both cases — was more controversial. Its assessment was presented in considerable scientific detail, including a full inventory of contents and analyses of the impact of leakage at different rates and depths, and of the physical and biological

pathways of toxins. Nevertheless, some scientists still complained that the oil company had failed adequately to explain how it arrived at its conclusions.

Shell still stands by its initial assessment and subsequent choice, which it says was based on reports from three independent consultants, as well as input from UK government scientists. Eric Faulds, project manager for the decommissioning of the Brent ►

Putting hazards under the microscope

London. Last October, the British government decided, in the wake of the public furore over the fate of the Brent Spar, to set up a group to reassess the various disposal options — and thus, hopefully, provide a firmer scientific basis for public discussion of its fate with possible implications for similar debates in the future.

The group, which is headed by John Shepherd, director of the Southampton Oceanography Institute, is now faced with the task of drawing together a diverse literature on the many scientific issues that bear on the environmental risk involved.

According to Shepherd, the group plans to draw as firm conclusions from

current scientific knowledge as it can. But it also intends to indicate areas of uncertainty, while it is using a series of natural analogues and scientific models which could give some idea of the context and the scale of the environmental hazard posed by the Brent Spar disposal.

Armed with a new inventory commissioned by Shell last summer from the Norwegian certification authority, Det Norske Veritas, the group is making an assessment of the toxic materials contained in the Brent Spar that could have found their way into organic material in general — and the food chain in particular — so as to help to identify how each is likely to interact with the environment.

The natural analogues and models that will be used to help assess the potential impact include the results of studies of sites of sludge disposal in the sea, as well as the seepage of natural methane from sediments on the continental shelf, to assist the understanding of biodegradation and the cycling of hydrocarbons.

The potential fate of heavy metals from the Brent Spar will be assessed using

studies of corrosion in shipwrecks, including the Titanic, which lies at about the same depth (more than 2,000 metres) as the originally planned disposal site for the Brent Spar. The analysis will also consider a photographic survey of containers holding low-level radioactive waste that have already been dumped in the north-

east Atlantic, as well as natural hydrothermal vents — areas of the seabed on mid-ocean ridges which force metal-rich solutions into the environment.

Despite the relatively small amount of PCBs on the Brent Spar, the group will also draw on a study of PCB pollution of the Hudson River in Canada,

where an electrical plant was found to have been releasing PCBs over a long period of time, extrapolating the results from shallow to deep waters.

Another assessment will be made of knowledge of the biota at the depths where Brent Spar was to have been dumped. At such depths, the amount of biomass found is between ten and 100 times lower than in shallow water. But, according to Shepherd, assessing biodiversity requires a "bold extrapolation" from known data; moreover, it is unclear whether a high or low biodiversity would best favour species survival.

The review group will issue its report this month. According to Shepherd, it is likely to conclude that, from the scientific point of view, there will be little environmental damage. The physical site of potential damage is small — the size of two football pitches. And fish stocks are unlikely to be affected, as some had claimed, as there are no biological or physical means by which toxic material can be carried up from such a depth to the layers closer to the surface that are fished commercially.

A. A.

►Spar, says that it was clearly the best option on the basis of all three of the major criteria used; even the environmental risks, says Faulds, are higher for the onshore disposal option because of the danger that the buoy might break up as it was being towed to shallow waters.

But the oil company reversed its decision when Greenpeace attacked its disposal plan, with protesters trailing the Brent Spar to its proposed watery grave, and hundreds of thousands of sympathetic motorists boycotting Shell, particularly in Germany, where service stations were firebombed.

Of course, there is no certainty that even the best scientific information could have defused the conflict. But the company now admits that its process for selecting deep-sea disposal should have been more transparent. And all parties might well have benefited from a more thorough and systematic risk assessment of the various options being studied.

Now, with the Brent Spar towed to a deep-water site in a Norwegian fiord, such an assessment is finally under way. An independent group of scientists, headed by John Shepherd of the Southampton Oceanography Centre, has been given the task by the UK Department of Trade and Industry.

The new study may lay to rest some of the disagreements about the potential environmental impact of the original deep-sea disposal plan. In particular, it is expected to conclude that Shell's own analysis was substantially accurate in its assessment of the impact.

But public statements by Greenpeace suggest that, whatever the outcome of the Shepherd review, its own preference for land disposal is unlikely to change. In particular, the environmental group claims that the impact of cumulative disposal of many such structures over very long time periods remains unresolved, and argues that the Brent Spar's materials should be recycled.

Last year's high drama is now over. But a decision about the fate of the Brent Spar still needs to be made. In October, Shell invited proposals from individuals and contractors for the disposal or recycling of the storage buoy. Later this month, it will announce a short list of six, and these will be further assessed not only on the usual criteria of safety, cost and environmental risk but also, as a result of last summer's events, public acceptability.

Meanwhile, Greenpeace's most successful argument is likely to be that no amount of scientific argument can offset a gut feeling that the sea should not be used as a toxic dumping ground under any circumstances. Shepherd himself acknowledged that it is not only scientific considerations that should be considered in reaching an eventual decision. "If people have an emotional response to pristine areas like Antarctica or the deep sea, and want them to remain unpolluted, it is not up to scientists to say this is irrational," he says. **Alison Abbott**

Brent Spar's risks: who says what

London. Throughout the Brent Spar affair, Shell has argued that its choice of sinking the oil storage buoy off northwest Scotland was based on a careful analysis of the environmental impact and health risks, using independent expertise. Documents made public by the company cite several reports that it commissioned to examine the Brent Spar disposal, including those produced by researchers at the University of Aberdeen.

But Greenpeace has disputed Shell's claim to have made a dispassionate assessment of the various options. It has argued that Shell chose the most convenient option, that its conclusions were questionable, and that the company was unduly secretive about the data it used. Despite subsequently acknowledging that faulty sampling methods had led to its own incorrect assessments of sludge on the buoy, Greenpeace continues to insist that its own arguments still rule out deep-water disposal.

Following Greenpeace's victory over Shell, the two sides have continued to cross swords over the disposal of the Brent Spar through a flurry of public statements — including allegations about the use of science in the affair. Two of the key themes have been:

Decision-making framework

Greenpeace questions Shell's decision to treat the disposal of such installations on a case-by-case basis. The group argues that this excludes consideration of the cumulative impact of future disposals, and that it therefore runs counter to the tide of international opinion, potentially establishing a precedent for dumping other wastes into the sea.

Greenpeace also points out that all signatories to the 1958 Geneva Convention on the Continental Shelf (including the United Kingdom) are committed to removing abandoned installations completely. It also says that, as the effects of dumping toxic material into the sea are unknown, Shell's position runs contrary to the "precautionary principle" enshrined in a number of international agreements to which the United Kingdom is also a signatory.

Greenpeace maintains, moreover, that Shell wrongly excluded options that, according to an analysis commissioned by Greenpeace, would allow 99 per cent of Brent Spar materials to be recycled or made harmless.

Shell responds by pointing out that Brent Spar is a one-off structure and therefore does not set any precedent. The company alleges that Greenpeace is conducting "single-issue campaigning" — that it is concerned only with the potential toxic effects of the structure on the ocean bed.

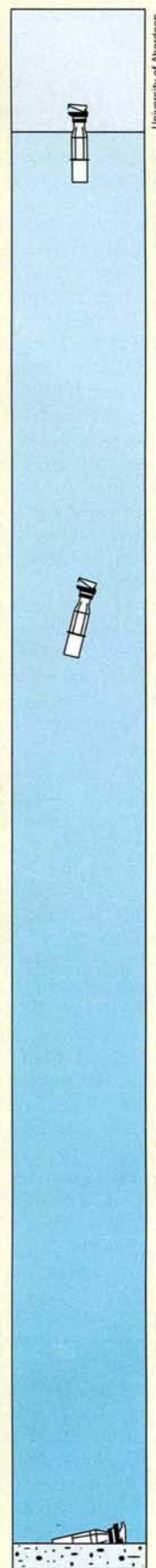
The oil company, on the other hand, argues that the impact of waste on one area of the environment cannot be considered in isolation from other areas — such as land and the atmosphere — nor can it be considered in isolation from the impact on human health, safety and cost. Furthermore, says Shell, Greenpeace "have not seen the full details of studies undertaken by numerous expert bodies", which it intends to publish in full.

Bad science

Both sides accuse the other of poor use of science. Greenpeace claims that Shell kept many scientific documents confidential for commercial reasons, that "there was no open, rigorously critical appraisal of the science" and that Shell had been selective in its choice of experts consulted.

Greenpeace asks, for example, why Shell did not take advice from the Scottish Association for Marine Science (SAMS) and the Natural History Museum, two bodies with long experience whose scientists criticized the Shell proposal once it was made public.

Shell has responded by saying that it now welcomes comments by SAMS scientists, and that it is "willing to support any rational forum in which they can be further discussed". But the company points out that "the points made by these scientists were all addressed by the UK government's licensing authority" and that, over a three-year period of study into the planned disposal, it had consulted many independent experts.



University of Aberdeen

Nature a back number?

SIR — Your leading article about the Engineering and Physical Sciences Research Council (EPSRC), “An agency that is flexible but flawed” (*Nature* 379, 753; 1996), hints at an important issue, namely the design of procedures by which a research support agency can decide financial allocations to its different subject areas; the advice offered, however, is shown by your own arguments to be inadequate and the evidence on which the solutions are based is far from the truth.

Three statements can be put to the test. First, the “United Kingdom’s physics community is having a rough time”. This ignores the fact that, within the EPSRC, physicists have a higher level of funding support than any other community. The comparison with chemistry (the next largest group) is shown in the figure. If facilities are included, the advantage to physics is even greater. Furthermore, responsive mode proposals, where the topic is the choice of the applicant, have a success rate in the physics programme of 53% (chemistry 27%) (EPSRC

abuse of the system. It will indeed be music to the ears of those promoting perpetual motion machines.

Research priorities must be established in a way that is effective and that can win the respect of the community. You complain that the “list of criteria... is dominated by factors relating to industrial potential”. You publish these criteria and readers can therefore assess the extent to which they are “troublesome”. The positive treatment of both chemistry and mathematics when exposed to such criteria should, however, provide reassurance about the flexibility of their impact.

In seeking solutions, you suggest that “standing disciplinary committees” are a regrettable absence from our prioritization procedures as we issue “signals that physicists are not doing as well as others in coming up with good ideas”. You yourself recognize, however, the precise weakness of such committees in this role when you ask “what learned society would sensibly start choosing winners and losers from among its membership?”

The system the EPSRC has in fact set in place must undergo refinement in the light of experience; it is, however, founded on three simple statements that would, I hope, find sympathy with the wider community:

(1) Project selection must be made by peer review. Your disingenuous reference to “colleges” conceals the fact that these are formal assemblies of peer reviewers nominated by the research community (*EPSRC Colleges*, and a *Guide to Peer Review*, October 1995). Our programme managers are in post to manage the peer review process and not to carry out peer review itself. The model of the US National Science Foundation has been expressly avoided in all UK research councils precisely because of the powers of selection that it places in the hands of single individuals, even if “drawn from their communities”.

(2) In the classical disciplines, including physics and engineering, the task of creative research thinking is best left to the individual scientist. The comprehensive strategic directives from the former ‘standing disciplinary committees’ (which also conducted the subsequent peer review) have been replaced by two-page ‘landscapes’ (*EPSRC Programme 1995–96*, March 1995) of priority indicators reflecting the thinking, for example, of the Government Foresight Exercise.

(3) Priority setting must be flexible. Under the present process, carried out by Council annually, decisions benefit from the comments of all interested parties on the published statement (already reprinted three times on demand) of research priorities (*EPSRC programme 1995–96*, March 1995). The accumulated comment is

reviewed by the Council’s two senior panels (*Newsline* January 1995) together with information from the Foresight Exercise and from programme managers on the current portfolios of the programme areas. The two panels then debate priorities and advise Council. You are entitled to propose other methods, but standing disciplinary committees have not shown themselves to be well fitted to the task.

The EPSRC has perhaps been “radical and innovative” and there will undoubtedly be strengths and weaknesses in the aftermath of the changes. With constructive comment (and this has been received) the weaknesses are being addressed. The benefits of the new system (to which you make scant allusion) are, however, surely significant. They include:

- higher success rates
- faster response times
- more open and participative peer review
- open and succinct priority statements
- absence of submission deadlines
- sympathy to interdisciplinary themes (absence of community ‘territories’)

A global dismissal of the new system prompted by concern that the physics community may find it difficult to defend its funding position is patronizing to physics colleagues. More seriously, it reflects a disappointing rejection of the “radical and innovative”, whatever its merits. *Nature* finds itself clutching instinctively to the habits of yesteryear.

Richard J. Brook

(Chief Executive)

Engineering and Physical

Sciences Research Council,

Polaris House, North Star Avenue,

Swindon SN2 1ET, UK

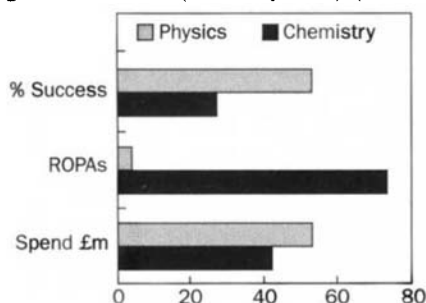
Knowing birds

SIR — Your review of Nan Dunbar’s inspiring edition of *The Birds* has drawn attention to Aristophanes’ contribution to ancient ornithology. But even earlier, the poet Pindar, in about 500 BC, reported that Zeus had practical knowledge of avian ethology. (Pindar’s account, since lost, was known to Strabo, the earliest geographer: see H. L. James (trans. and ed.), *The Geography of Strabo*, Heinemann, 1923.)

Zeus wanted to find the centre of the Earth, so he started two eagles of equal speed at the same moment, one from the eastern edge of the world, one from the western. They met at Delphi, and thus his putative hypothesis, presumably conceived on Olympus, was not refuted. The omphalos marks the spot where they fell, stunned by their collision. If this account is accurate, Zeus should, at last, be given credit for performing the first scientific experiment.

John Godfrey

41 Lawford Road, London NW5 2LG, UK



‘% Success’ is for responsive mode proposals, by programme; ‘ROPAs’ is the number of proposals in 1995, by programme; ‘Spend £m’ is spending by academic department (1994–95), covering grants and studentships.

letter to vice-chancellors, 4 October 1995). International comparisons suggest that this hardly corresponds to a “rough time”.

Second, “the EPSRC is turning the screw”. Even with allowance for editorial hyperbole, this seems an excessive reaction to the cut of approximately 3% to be made in the physics programme before the next prioritization exercise.

Third, “a relative lack of involvement with industry has counted against the physics community”. The figure compares the actual lack of involvement as measured by ROPA applications, where £25,000 of industrial cash is a requirement for entry into the scheme. “Relative” or not, this level of involvement is surely an item for legitimate consideration.

You say that “if these areas are not suited to attracting funds from elsewhere, that is all the more reason why the EPSRC should reinforce its commitment to them”: that would open the door to alarmingly wide

Endangered Spanish science

SIR — We should like to respond to Javier Garcia-Guinea (*Nature* 379, 109; 1996). He says there are two kind of scientists in Spain, those who publish in journals ranked in the *Science Citation Index (SCI)* and what he calls “parallel scientists”. He complains that the latter are denied grants or receive less funding and respect than their counterparts. And he says that they “... publish in serious Spanish journals that are not in the *SCI* lists. They often know French and German (rather than English) and carry out other activities such as teaching, writing books, organizing meetings or running laboratories.”

Most of the “serious” Spanish journals are not refereed journals in the accepted sense; the publisher, editorial board, referees and authors belong to the same small group of people. What makes a good journal is not whether it is written in English or Spanish but that publication is restricted to high quality papers.

As far as the language itself is concerned, for all the disciplines Garcia mentions, the most important papers and the leading journals are published in English. So how can researchers who cannot understand English know what is happening in their fields? And how can they communicate their exciting findings to the international scientific com-

munity? Science in the twentieth century consists not only of gathering exciting data or doing good experiments, but sharing the results in a common effort. We understand the deficiencies in the Spanish educational system, which only lately has been able to acknowledge the importance of English as the official language of science; true ignorance comes not from not knowing something but from not recognizing the need to learn.

Finally, the *SCI* scientists also teach (although perhaps less than the parallel ones), write books or book chapters (most of them in English, internationally distributed), organize international meetings and run laboratories (which are increasingly joined by researchers from other countries).

In Spain, as in any other country, there are two kinds of scientists, bad and good. Many of us are training abroad, trying to do quality research and facing now the uncertainties of an increasingly saturated job market. Many of us have tried to join Spanish institutions only to find the doors closed because we publish in *SCI* journals but do not have teaching experience. (The undergraduate and graduate training that we supervise counts for nothing.) Many of us know nobody among the parallel scientists who still govern many Spanish institutions,

and thus need to live with our heart in one place and our brains and hands in another. For us, the hope that mediocrity will be banished from Spanish science and that quality criteria will direct scientific policies is the only thing left.

Rosa Carballada

Cristina Hernández-Munain

Carlos Suñe-Negre

*Duke University Medical Center,
Durham, North Carolina 27710, USA*

SIR — Recent correspondents have drawn attention of the problems of Spanish scientists and the uncertain future. This is a sad paradox: between 1982 and 1992, “the rise in scientific production in Spain has been disproportionately greater than in the rest of the world” (F. J. Ayala, *Política Científica*, 43, 5–12; 1995). This increase has been largely the work of young scientists with postdoctoral training in foreign institutions. They are now returning to Spanish institutes under 3-year contracts to carry out particular tasks within continuing projects. Some 600 young scientists now have these 3-year contracts. They may account for 20–30 per cent of researchers in some flagship institutes of the Spanish Council for Scientific Research (CSIC).

Because the first opportunity to hire postdocs arose in 1992, some of these highly qualified researchers have now finished their contracts and are facing an uncertain



future. Some of them have chosen to go abroad again as postdocs, others have abandoned research after almost a decade, and many are joining the ranks of the jobless. The reason? There is a shortage of research posts. In the last opening for CSIC faculty, only 50 staff positions were available for the whole of Spain. That has been the situation for the past 3 years. Between 1982 and 1992, the amount invested in science increased from 0.48 to 0.9 per cent of Spain's gross domestic product (GDP); during the same period, the SCI impact factors of papers from Spain increased from 1.2 to 1.78.

Now, however, the same scientific authorities have failed to follow through. For most of us, these contracts were considered a springboard to an independent position. In the event, the conditions under which we were taken on preclude us from working on any other project. We can neither submit proposals for funding to national boards (such as the Interministry Committee of Science and Technology; CICYT), nor to any other institutions. So our hands are tied. Only staff scientists can apply for funds or be principal researchers on a project. Moreover, there are no tenure-track positions available. Funds for science need to grow to at least 2 per cent of GDP (as in other European Union countries) with a corresponding increase in the number of researchers (currently 1.4 per 1,000, compared with 1.9 per 1,000 else-

where in the European Union).

The future of Spanish science depends on a proper solution to the problems faced by these returned scientists, who can help to obtain the benefits expected from the past decade's effort and obtain funds in fair competition for international projects with other European colleagues and institutions.

Juan F. Beltrán

Juan J. Negro

José L. Yela

José A. Mateo

*Estación Biológica de Doñana, CSIC,
Apdo 1056,
41080 Sevilla, Spain*

Study in Germany

SIR — A recent leading article (*Nature* 378, 755; 1995) mentions a European Commission (EC) survey of young European scientists who applied for fellowships to train in another European Union country under the EC's Human Capital and Mobility (HCM) programme: only 10 per cent out of 6,000 applied to do research in Germany. You interpret these figures as a sign of "the unpopularity of Germany as a place for young foreign scientists".

I should like to point out that Germany, because it has a national sponsoring scheme to attract young foreign scientists, is less dependent on the new European pro-

gramme than other countries in Europe. Under existing programmes of the Alexander von Humboldt Foundation, the German Academic Exchange Service, the Helmholtz Association of National Research Centres and the Max Planck Society, more than 1,500 young foreign scientists and scholars from other European Union countries came to German research establishments in 1994. The Humboldt Foundation alone registered 40 per cent more applications than ten years ago. The increase in real sponsorship figures for the Max Planck Society was 80 per cent in 10 years.

Young foreign scientists are most welcome in Germany. English is a current language in many institutes and laboratories. German language courses are available in all national programmes. Many guest researchers have told us that German is not a difficult language to learn. Foreign guests are as a rule exempted from the burden of bureaucracy in universities. We are not aware of any hostility towards foreigners in universities. On the contrary, cooperation with foreign guests is seen as an enrichment for national research. We are grateful, therefore, for the positive conclusion to your leading article.

Manfred Osten

(Secretary General)

*Alexander von Humboldt Foundation,
Jean-Paul-Str. 12,
D-53173 Bonn, Germany*

If you've ever developed a purification strategy by yourself, you know there is a lot to consider. Creating media screening schemes; designing buffer preparation routines; selecting which columns to use—even listing the tasks takes careful consideration. But now there's a better way of working.

Introducing ÄKTAexplorer

Turn ÄKTA[™]explorer on, choose a protocol, check running parameters and press start. ÄKTA-explorer does the rest.

Designed for all chromatographic techniques, this new purification system offers you pre-set protocols for every major purification task—including method scouting and media

screening. You'll save time as the system automatically recommends the best columns for your runs. You get fast pH screening as it automatically prepares your buffers from stock solutions. The moment you turn ÄKTAexplorer on, you're presented with a direct path to full-scale purification.

That path is UNICORN[®]—ÄKTAexplorer's control software. Your scale-up is simplified as the system's software can transfer and implement your methods on purification systems at all scales.

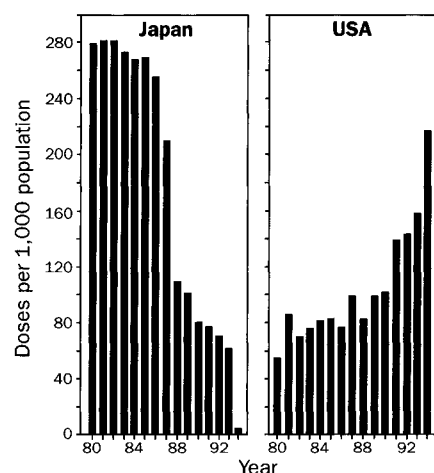
Of course, these are just a few of ÄKTAexplorer's features. So call us at 1 (800) 526 3593 from North America, +81 (0)3 3492 6949 from Japan or +46 (0)18 16 50 11 from the rest of the world (or meet us on the Internet at <http://www.biotech.pharmacia.se/akta.htm>).



Pharmacia Biotech
Uppsala, Sweden. (And the rest of the world)

Japan lagging in influenza jabs

SIR — Influenza epidemics have high costs in morbidity, mortality and medical care. Vaccination of high-risk groups, such as the elderly or those with chronic illness, is the most effective measure for reducing the impact of the disease¹. In the United States, the number of influenza immunizations has more than doubled in the past 10 years, and in 1994 coverage reached approximately 60% among people aged 65 and over^{2,3}. Furthermore, the number of doses of vaccine distributed per 1,000 population more than doubled between 1990 and 1994 alone, as shown in the figure. The trend is similar to that seen in other developed countries, although several-fold differences are still seen in per capita vaccine distribution⁴. In contrast, influenza vaccine distribu-



Influenza vaccine distribution in Japan and the United States, 1980–94, expressed as the number of doses distributed per 1,000 population. Japan's figure in 1994 is shown as the doses produced per 1,000 population, because accurate information on vaccine distribution is not available for that year.

tion in Japan has shown a marked decrease over the same period.

Influenza appeared on the list of targeted diseases in Japan from 1976 until June 1994. This policy was intended to control influenza epidemics in the entire community through suppressing transmission in schools⁵. Although the vaccination rate among schoolchildren was about 80% up to the early 1980s, it showed a steep decline thereafter, to 18% in 1992. This decline was due to poorly designed studies that alleged that the vaccine had little if any efficacy. These reports create scepticism about vaccination, not only among the general public but also among many medical professionals⁶.

Influenza is considered a minor illness in Japan. There is no national recommendation of target groups for active immunization, nor any system for reimbursement. In this rapidly ageing society, very few older people volunteer for inoculation, and when

they do they must pay approximately US\$50 to be vaccinated. Furthermore, Japanese pharmaceutical companies are decreasing or discontinuing their production of influenza vaccine. This contrasts strikingly with the United States where, since 1993, the federal government's Medicare programme has provided reimbursement for the cost of influenza vaccine and its administration², and where the manufacturers have greatly expanded their production.

Many countries are developing comprehensive strategies to cope with the next influenza pandemic. Given recent experience, it seems unlikely that attitudes towards influenza and influenza vaccination in Japan will change in the near future. As a result, Japan may be ill prepared to join other countries in international efforts to confront the next pandemic.

Yoshio Hirota

Department of Public Health,
Faculty of Medicine 60,
Kyushu University,
Fukuoka 812-82, Japan

David S. Fedson

Pasteur Mérieux MSD,
3 Place Antonin Perrin,
F-69007 Lyon, France

Masaro Kaji

Kurume University,
Kurume, Fukuoka 830, Japan

1. C.D.C. Morbid. Mortal. Wkly Rep. **43** (RR-9), 1–13 (1994).
2. Marwick, C. J. *Am. med. Ass.* **272**, 1640–1642 (1994).
3. Marwick, C. J. *Am. med. Ass.* **273**, 765 (1995).
4. Fedson, D. S. *et al. Vaccine* **13**, 623–627 (1995).
5. Dowdle, W. R. *et al. J. infect. Dis.* **141**, 258–264 (1980).
6. Hirota, Y. *et al. Lancet* **344**, 408–409 (1994).

Distrust of science

SIR — Scientists complaining of the poor and declining image of themselves and of science itself have, following John Maddox's Commentary (*Nature* **378**, 435–437; 1995) offer various remedies that scientists might exploit to improve the situation (*Nature* **379**, 292; 1996). Each of these fails to take into account a much more deep-rooted cause of the perceived malaise; a sea-change in the attitude of scientists themselves to the nature of science.

For a long period, from the mid-nineteenth century probably until the early 1960s, science was driven on a great inductive wave and characterized by widespread acceptance of the notion that man, from observations of his surroundings, could derive universal laws, thereby conferring a degree of authority on science that had previously been the preserve of religion; and this at a time when religious authority was anyway in decline.

But for some time now, the perceived authority of inductive logic as the means of

scientific advance has been in retreat: since *The Logic of Scientific Discovery*, scientists themselves have increasingly been encouraged to a more sceptical view and much of the scientific community now accepts many of Karl Popper's ideas on the deductive method of critically testing scientific ideas. This shift within the scientific community from a confident certainty to sceptical analysis of its theories is not simply a reaction to the well publicized failures of science; it is based on logically supported epistemological arguments, positively asserted by many scientists.

Much of the thrust of the discussion as to why the public distrusts science has centred on what is perceived to be an increasing chasm between scientists and the wider community. However, even though scientists may complain (with some justification) that many aspects of scientific endeavour are undervalued by the general public, they cannot ignore the fact that much public distrust is an accurate reflection of current scientific philosophy. It is not immediately obvious why the principle of healthy scepticism of scientific truths should be perceived as laudable within the scientific community but as unfortunate within the population at large.

Christopher B. Johnson

Department of Botany,
University of Reading,
Whiteknights, PO Box 221,
Reading RG6 6AS, UK

Pensionable age?

SIR — I am not surprised to hear Russian politicians criticizing the Soros scholarships for science and education. What does puzzle me is a similar reaction of scientists such as V. I. Rigin (*Nature* **378**, 432; 1995). He believes that the scholarships awarded to Russian professors "exert a negative influence" on our educational system, as they go to old (meaning inefficient) hands and thus prevent them from being duly ousted. (I should perhaps say that I do not hold a Soros scholarship.) The same reasoning can apparently be applied to any outside financial help, whether regular or irregular.

No correlation between the number of scholarships and age of recipients is given, which alone makes Rigin's logic flimsy. More important is that there is no universal correlation between the age of a teacher and the quality of teaching. Some tribes have been known to regulate the average age of the population by withholding food from those considered a burden of the community. At what age does Rigin consider a professor ready to be disposed of?

Yuri Berlin

Shemyakin-Ovchinnikov Institute
of Bioorganic Chemistry,
Russian Academy of Sciences,
Moscow 117871, Russia

A crack in the Standard Model?

Frank Wilczek

The Standard Model of particle physics has it that quarks are fundamental particles. But an experiment that echoes two past revolutions in our understanding of matter might hint instead at an even smaller level of structure.

THE Standard Model of particle physics describes how the different fundamental particles of matter interact with each other (see box overleaf). Since it assumed its current form in the early 1970s it has gone from triumph to triumph. Ever more detailed calculations have been confronted with ever more refined experiments, and in every case where unambiguous comparison has been possible, satisfactory agreement has been found. Discrepant observations have been reported on occasion, but so far none has stood the test of time. Now the CDF collaboration at Fermilab, which last year helped complete the edifice of the Standard Model with its discovery of the long-anticipated top quark, has come forth with new observations¹ that appear to challenge a central component of the model, the theory of quantum chromodynamics (QCD).

It is both significant and ironic that the apparent challenge comes as a sort of coda to one of the most awesome and impressive successes of QCD (Fig. 1). The theory accounts in quantitative detail for the production of distinctive collimated 'jets' of particles emerging at large energies and angles from proton-antiproton collisions. Such jets are the nearest thing we have to a tangible embodiment of quarks and gluons. It is notorious that, though they are the fundamental building blocks from which we construct our description of matter in the Standard Model, quarks and gluons cannot exist as free isolated particles. Yet the jets of ordinary particles into which high-energy quarks and gluons eventually materialize accurately reflect their original energy and momentum. QCD predicts in considerable detail the probability that such jets are emitted from collisions of the sort studied in the CDF experiment, as a function of their energy and angle.

As Fig. 1 illustrates, there is truly remarkable agreement between theory and experiment over a considerable energy range — a range over which the cross-section changes by several orders of magnitude. It is primarily in contrast to this remarkable concordance that a residual discrepancy at the highest energies (Fig. 2) stands out. The actual number of events involved in the discrepancy — a few hundred out of trillions of observed collisions, thinning out to a handful for the highest energy jets — is relatively

small, and individually they do not display any very obviously anomalous character.

What could it mean? For a naive physicist, the first encounter with the CDF phenomenon might trigger a sense of *déjà vu*. Excess scattering events at high energy have twice before, in celebrated experiments, pointed towards the existence of unexpected substructures. Rutherford interpreted the classic Geiger-Marsden experiment (which he had suggested) as indicating the existence of a small, hard atomic nucleus. And the deeply inelastic

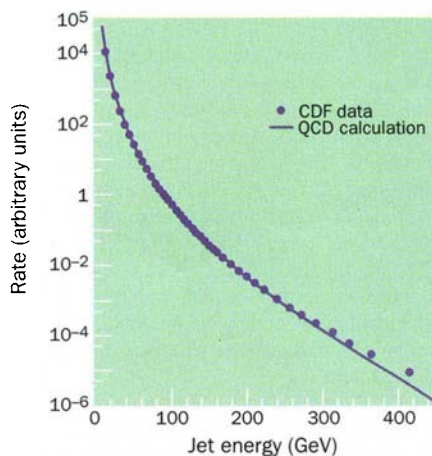


FIG. 1 The agreement between QCD theory and experiment reported by the CDF group. The rate of jet events is plotted against collision energy. A tiny discrepancy is visible at the highest energies. (Adapted from ref. 1.)

scattering experiments of Friedman, Kendall and Taylor gave decisive evidence for the existence of substructure — quarks and gluons — within protons. Could the CDF phenomenon be indicating that quarks and gluons, in turn, are composed of more elementary bits?

In considering such questions more deeply, it is important to keep in mind a few profound peculiarities of QCD. First and foremost among these is that QCD is a precisely defined, principled theory. It is derived from a fundamental assumption about the symmetry of the world (gauge invariance) that cannot be fudged. You must take it or leave it. There is, in principle, no room for ambiguity about the consequences of the theory.

In practice, however, the situation is more complex. The equations of QCD are not easy to solve. Relatively simple yet accurate calculations of basic high-energy

scattering processes among quarks and gluons are possible, due to the property of asymptotic freedom (at very high energies they behave almost as free single particles). But at present our ability to compute low-energy interactions, including those responsible for binding the quarks together to make protons, is much more limited because of the many-body nature of the interactions and all the virtual particles that affect the result.

In calculating what QCD predicts for CDF's proton-antiproton collisions, inevitably the gold and the dross come mixed. The important interactions involving quarks and gluons can be calculated accurately, but the distribution of these particles inside the original protons and antiprotons cannot. Instead these distributions must be inferred by constructing a consistent fit to a mass of data, covering a wide range of collision energies and several distinct types of experiment (see for example ref. 2 for a review). The fit is for the most part highly overconstrained. Over time, as successful predictions based on it have become routine, particle physicists have come to rely on the ability of QCD to forecast accurately the bulk of what they observe at accelerators. Figure 1, apart from the barely visible discrepancy at the highest energies, is broadly representative of many similar comparisons of experimental results to sharp QCD predictions.

Nevertheless there are significant uncertainties in the quark, and especially the gluon, distributions. The interpretation of the CDF results which would least disturb our dogmatic slumber is that the relevant QCD calculations have been performed incorrectly, or that their uncertainties have been underestimated. One can expect that these issues will be intensely scrutinized in coming months.

Pending the outcome of such scrutiny, it is fun to entertain the possibility that the CDF observations might represent fundamentally new physics. What could that be? Let me mention some possibilities that leap into the minds of theorists, and have already received some attention.

As I have already emphasized, it seems very difficult to modify the core of QCD — the gauge symmetry that dictates the nature of the interactions among quarks and gluons — without destroying the theory altogether. Given the overwhelming success of the theory, physicists will

Getting up to Standard Jargon

THE Standard Model comes equipped with a Standard Jargon. Despite appearances, this jargon was not expressly designed to bewilder and intimidate the uninitiated. A crash mini-course in remedial jargon follows.

The Standard Model has two components. One is a theory of the strong interaction, called quantum chromodynamics or QCD. The strong interaction is traditionally said to be the interaction responsible for nuclear forces. That is true, but a more modern point of view is that QCD is a theory of interactions among quarks and gluons (out of which atomic nuclei are made). The other component is a theory that gives a unified description of weak and electromagnetic interactions — mercifully this has no special name, and is called simply electroweak theory.

The physical concepts used in the Standard Model are generalizations of

concepts familiar in electrodynamics. The quarks and gluons are said to be 'coloured' in various combinations of red, white and blue. What this means is that they carry three distinct types of charge designated by those chromatic names. In electrodynamics, the ordinary photon responds to electric charge. In quantum chromodynamics, the various gluons (there are eight of them) respond to colour charges. An interesting and very important new feature of QCD, as compared with electrodynamics, is that the gluons themselves carry colour charges, unlike the photon which is electrically neutral. Although the equations of quantum chromodynamics are very similar to those of quantum electrodynamics, because of this new feature they are much more difficult to solve and lead to very different physical behaviours.

While gluons are analogues of photons, quarks are analogues of electrons.

They carry not only colour charges but also ordinary electric charge. A crude but not entirely wrong model of a proton is that it consists of three quarks held together by constant exchange of gluons.

In the accompanying article you will also come upon the phrase "vector bosons similar to the Z boson, but heavier". The Z boson is a particle mathematically very similar to the photon, but differing in two crucial respects. First, it is massive (the photon has no mass), weighing close to a hundred times as much as a proton. This drastically affects its physical properties, making the Z boson unstable and rendering the forces it mediates very short-range. Second, it responds to an abstract charge that is not the same as electric charge or any colour charge. First observed directly in the early 1980s, the Z boson is an important ingredient in the electroweak sector of the Standard Model. F. W.

consider this option only under extreme duress. Much more palatable is the possibility of leaving the core of QCD intact, while adding some extra interactions to account for phenomena that don't fit.

One possibility is that the quarks have some sort of super-strong interaction in addition to the strong interaction described by QCD. If all the particles mediating this interaction were heavy, then their influence, although ultimately dominant, would only become visible at high energy — a pattern broadly consistent with the CDF observations. Unfortunately there are no good reasons to welcome such an interaction, and considerable reasons to fear it, for it would tend to render coincidental the marvellous unification of forces³ that occurs in simple extensions of the Standard Model.

Another widely anticipated possibility is the existence of new vector bosons similar

to the Z boson, but heavier. If such Z' bosons were actually being produced in the CDF experiment, they should decay into pairs of jets whose total effective mass would be the Z' mass. The observations, however, do not indicate the existence of such a favoured effective mass for jet pairs. This is not the end of the story, though, because the exchange of virtual Z' bosons could affect the probability of producing high energy jets, even at energies insufficient to produce the bosons themselves.

An extra incentive to consider such possibilities comes from the existence of another small but long-standing discrepancy between the Standard Model and experiment. The rate of Z decay into bottom-antibottom quark pairs appears to be a few per cent higher, and into charm-anticharm pairs a few per cent lower, than the electroweak component

of the Standard Model predicts⁴. A Z' boson, if it existed, could mix with the conventional Z boson and alter its properties. It is possible to find combinations of parameters^{5,6} such that both of these electroweak anomalies are accommodated, together with the QCD anomaly found by CDF. There is, however, a weighty theoretical objection to this line of thought. In order to generate an effect as large as that claimed by CDF, the interaction strengths of the Z' must be quite large. When extrapolated to only slightly higher energies, they actually go to infinity, presumably indicating

that the theory is not self-consistent.

Another widely anticipated and extremely attractive possibility is the existence of a whole new world of particles, superpartners of the known particles, a realization of the idea of supersymmetry (ref. 7 and references therein). Again, if CDF were actually producing the superpartners, one would expect to see effects that are not seen — in this case, a gross enhancement in the number of jets and momentum imbalances due to the escape of unobserved neutral particles. So one must appeal, as for the Z' boson, to interference effects from virtual particles. Unfortunately, despite the abundance of new particles, explicit calculation shows that the effect of supersymmetry below production threshold is too small to explain the CDF observations⁸.

So no ready explanation appears compelling. If the CDF phenomenon really does represent fundamentally new physics, it probably takes a form not yet widely contemplated. □

Frank Wilczek is in the School of Natural Sciences, Institute for Advanced Study, Olden Lane, Princeton, New Jersey 08540, USA.

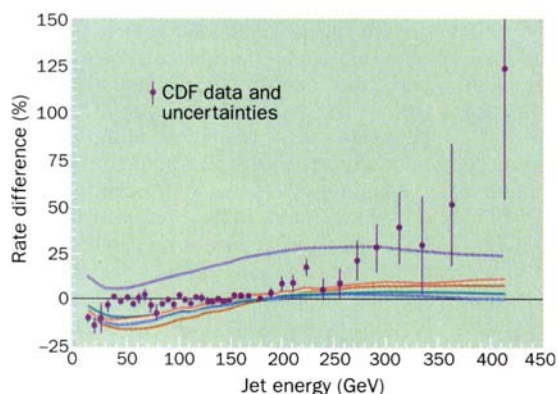


FIG. 2 The relative discrepancy between QCD theory and experiment reported by the CDF group. The experimental values are compared with a number of different calculations of QCD (each using different assumptions about how quarks and gluons behave inside protons) and normalized to one particular calculation. (Adapted from ref. 1.)

1. Abe, F. et al. *Phys. Rev. Lett.* (submitted).
2. Zerwas, P. & Kastrup, H. (eds) *QCD — 20 Years Later* (World Scientific, Singapore, 1993).
3. Dimopoulos, S., Raby, S. & Wilczek, F. *Physics Today* **44**, 25–33 (1991).
4. The LEP Collaborations ALEPH, DELPHI, L3, OPAL, and the LEP Electroweak Working Group CERN-PPE/95-172 (CERN, Geneva, 1995).
5. Altarelli, G., Di Bartolomeo, N., Feruglio, F., Gatto, R. & Mangano, M. CERN-TH/96-20 (CERN, Geneva, 1996). FTP from hep-ph/9601324.
6. Chiappetta, P., Laysac, J., Renard, F. & Verzeqnas, C. (1996). FTP from hep-ph/9601306.
7. Kane, G., Kolda, C., Roszkowski, L. & Wells, J. *Phys. Rev. D* **49**, 6173–6210 (1994).
8. Kraus, P. & Wilczek, F. (1996). FTP from hep-ph/9601279.

Big rigs in blood coagulation

J. Fernando Bazan

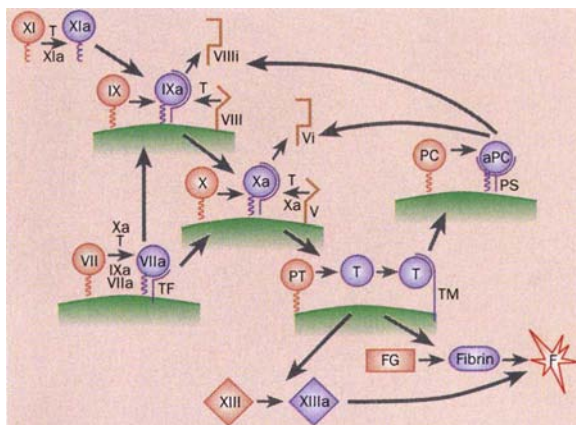
CUTS bleed. But usually not for long, because elaborate mechanisms have evolved to limit the loss of blood after vascular damage and to speed healing of the wound by inflammation and repair responses¹. A long-standing puzzle has been how, from the interaction between two of the players in the blood-clotting cascade (tissue factor and coagulation factor VIIa), the clotting process is triggered. Much of that puzzle is solved by the paper by Banner and colleagues on page 41 of this issue², which describes the structure of the complex between the two, and which bears more generally on the assembly and function of protein complexes in blood clotting.

Two initial events in clotting are the formation of a cellular plug from the adhesion and aggregation of circulating platelets, and the generation of an insoluble fibrin network. The molecular and biochemical basis of this latter defence — blood coagulation — frustrated researchers for years until, in 1964, an elegant outline of an 'enzyme cascade'³ and a 'waterfall sequence'⁴ crystallized the notion that a repeating series of proteolytic conversions led to the production of fibrin from plasma fibrinogen.

The key principles of this conceptual framework for the coagulation cascade — amplification, regulated activation and protease components — remain in the modern scheme (see figure), with the addition of two critical adjuncts: the dependence of many of the activation steps on phospholipid membranes and on non-enzymatic protein cofactors⁵. Banner and colleagues' remarkable crystal structure, which provides the first three-dimensional view of an activated coagulation enzyme bound to its membrane cofactor, physically embodies these myriad, intersecting demands of the coagulation cascade.

The amino-acid sequences of the various factors, cofactors and regulatory proteins at work in the cascade show recurring modular themes in their structural design⁶. The enzyme factors typically feature a carboxy-terminal, trypsin-like, serine protease domain and an amino-terminal module rich in Ca²⁺-binding, γ -carboxyglutamic acid (Gla) residues; these domains sandwich two epidermal growth factor (EGF)-like modules in the case of factors (F) VII, IX, X and protein C, or two kringle motifs in prothrombin^{5,6}. There is a greater diversity of forms among the cofactor chains, but all retain an ability to associate with

phospholipids: FV, FVIII and protein S through Gla domains, and tissue factor (TF) and thrombomodulin through transmembrane segments. The sequential conversion of active protease factors from inactive zymogens occurs through specific proteolytic cleavages (see figure) that simply sever catalytic or 'heavy' domains from



The blood coagulation cascade, beginning with the interaction between FVIIa and tissue factor (TF), emphasizing the assembly of complexes between membrane-bound factors and cofactors. The last step in the cascade, the crosslinking of fibrin into clots, is carried out by the transglutaminase factor XIII. Protein C (a close relative of factors VII, IX and X) has anticoagulant effects by inactivating (i) factors V and VIII. Most of the molecules display phospholipid-association domains (drawn as tails), or transmembrane segments; many of these reactions occur in the presence of Ca²⁺. Inactive, zymogen forms of factors and cofactors are depicted as red circles or diamonds; proteolytic maturations are indicated by arrows; active enzyme forms are denoted by 'a'. T, thrombin; PT, prothrombin; TM, thrombomodulin; PC, protein C; PS, protein S; FG, fibrinogen. For a detailed review see ref. 5.

their amino-terminal 'light' modules; these chain segments remain attached by single disulphide links. The component folds of the FVIIa domains in Banner and colleagues' complex were expected from earlier X-ray and NMR studies of related module fragments, and the kindred protease structures of thrombin, FXa and FIXa (refs 7–9). In turn, the TF fold complexed with FVIIa seems to be closely similar to the unliganded structure of soluble TF^{10,11}, a tandem set of immunoglobulin-like β -barrels homologous to class 2 or interferon-like haematopoietic cytokine receptors^{12,13}.

The chief impact of the new structure lies instead in the contoured embrace of FVIIa and TF protein folds, and how this peculiar engagement produces a highly active molecular machine that attacks FX and FIX substrates and triggers the main pathway of the coagulation cascade (see figure). Shaped like a question mark, the FVIIa half of the complex places the catalytic domain on the top rim and the light chain along the curved stem, effectively forming a hollow that grasps the top

(amino-terminal) β -barrel of TF; the bottom TF domain stands alongside a FVIIa stem that seems to be punctuated at its end by a cluster of bound Ca²⁺ ions in the Gla domain.

The dimensions, geometry and extended packing mode of the crystal complex were predicted quite accurately by Ashton *et al.*¹⁴ from an X-ray and neutron scattering study of FVIIa bound to TF in solution. But if we were to infer the attitude of bound FVIIa from available coagulation factor templates, we might not perform as well: although the backbones of the protease and proximal (EGF-2) domain are superimposable between FVIIa and equivalent parts of the FXa and FIXa structures, the resulting orientation of the EGF-1 and Gla modules turns sharply askew in the FIXa stem⁹, and now runs parallel to the deduced phospholipid plane in the FVIIa:TF complex. This conformational divergence maps exactly to a flexible hinge between EGF-1 and EGF-2 domains in FVIIa and FIXa.

The intermolecular contacts between FVIIa and TF are extensive: about 1,800 Å² of total surface area is buried in the complex, most of which is due to the lengthwise packing of the FVIIa light chain against TF; similar values are observed in the dockings of other large, heterodimeric proteins¹⁵. The light chain interface paints a swath of residues in TF that are largely located along the edge β -strands on one side of the molecule; by contrast, the FVIIa catalytic domain perches on the top-facing β -sheet of the amino-terminal TF domain. This binding pattern in TF is very different to the loop-rich interactions of haematopoietic receptors with cytokine ligands.

Mutant forms of TF that are impaired in cofactor function segregate into molecules that are either affected in binding to

1. Ratnoff, O. D. *Perspect. Biol. Med.* **31**, 4–33 (1987).
2. Banner, D. W. *et al.* *Nature* **380**, 41–46 (1996).
3. MacFarlane, R. G. *Nature* **202**, 498–499 (1964).
4. Davie, E. W. & Ratnoff, O. D. *Science* **145**, 1310–1312 (1964).
5. Davie, E. W. *Thromb. Haemostasis* **74**, 1–6 (1995).
6. Patthy, L. *Semin. Thromb. Haemostasis* **16**, 245–259 (1990).
7. Stubbs, M. T. & Bode, W. *Curr. Opin. struct. Biol.* **4**, 823–832 (1994).
8. Padmanabhan, K. *et al.* *J. molec. Biol.* **232**, 947–966 (1993).
9. Brandstetter, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **92**, 9796–9800 (1995).
10. Muller, Y. A., Ullrich, M. H. & de Vos, A. M. *J. molec. Biol.* **256**, 144–159 (1996).
11. Martin, D. M. A., Boys, C. W. G. & Ruf, W. *FASEB J.* **9**, 852–859 (1995).
12. Walter, M. R. *et al.* *Nature* **376**, 230–235 (1995).
13. Bazan, J. F. *Nature* **376**, 217–218 (1995).
14. Ashton, A. W. *et al.* *FEBS Lett.* **374**, 141–146 (1995).
15. Jones, S. & Thornton, J. M. *Proc. natn. Acad. Sci. U.S.A.* **93**, 13–20 (1996).
16. Ruf, W. *et al.* *Biochemistry* **35**, 6310–6315 (1995).
17. Esmen, C. T. *FASEB J.* **9**, 946–955 (1995).

FVIIa, or have a reduced ability, as FVII:TF complexes, to activate FX (refs 2, 10, 11). Most of the binding mutations appear to map to the top half of the light-chain recognition swath that marks the EGF-1 and EGF-2 docking sites in TF. By mapping these data onto the TF scaffold^{10,11}, last year Ruf *et al.*¹⁶ managed to produce a very credible picture of the surface features of TF bound by FVIIa. Clustering instead near the bottom of the carboxy-terminal TF domain^{10,11}, at the level of the FVIIa Gla domain², the activation-deficient mutations lie some 75 Å beneath the gaping maw of the protease active site. This latter epitope probably defines a membrane-proximal site of TF interaction with FX that optimally poses this substrate for proteolytic cleavage by bound FVIIa.

The observed increase¹¹ of FVIIa catalytic activity on binding TF may be a response to several restraints that are imposed by the structure of the complex. At a minimum, TF serves as a rigid scaffold and secure membrane tether for FVIIa that effectively reduces the mobility of the factor chain, perhaps locks the protease into an open or 'active' conformation, and facilitates the assembly of ternary complexes with equally bulky substrate mol-

ecules^{2,9,11}; in accord with this view, free FVIIa in the presence of Ca²⁺ and phospholipids remains only weakly active¹¹. More intriguingly, TF may 'lift' the FVIIa active site to within a narrow shell or zone 60–70 Å above the cell membrane where other coagulation factors or cofactor complexes convergently display their scissile bonds, or catalytic machinery. This view is consistent with fluorescence energy transfer and electron microscopy experiments, which show a remarkable uniformity in the extramembrane projection of the active sites of FIXa:FVIIIa, FX:FVa and thrombin:thrombomodulin complexes along with FVIIa:TF (refs 2, 9, 17).

A functional imperative in the coagulation cascade may be the evolutionary maintenance of these tall molecular 'rigs' as raised catalytic platforms for the major activation steps, cushioned by distinct constellations of spacer modules and membrane anchors⁹. In the absence of critical cofactors, clotting proteases may not reach the necessary heights or field the right tools to do the job. □

J. Fernando Bazan is in the Department of Molecular Biology, DNAX Research Institute, 901 California Avenue, Palo Alto, California 94304-1104, USA. bazan@dnax.org

SOLAR SYSTEM

Bombarding Mars lately

Clark R. Chapman

A MOST unusual rock, picked up from the Antarctic ice over a decade ago, may tell us something profound about Mars — and about the first billion years of Solar System history. At first the rock, called ALH84001, was thought to be a normal meteorite. But a few years ago, studies¹ revealed textural features and oxygen isotope data proving that it is one of a clan of a dozen meteorites (called SNCs) believed to be from Mars². Now, on page 57 of this issue³, Ash, Knott and Turner report that ALH84001 is 4.0 billion years (Gyr) old, to within 0.1 Gyr — far older than the other martian meteorites, all of which are younger than 1.3 Gyr. It is only one rock, but ALH84001 may be the tip of an iceberg.

At issue is the nature of an epoch in planetary history termed the late heavy bombardment, first identified by analysis of the ages of Moon rocks⁴. The Moon was intensely bombarded from 3.8 until 4.0 Gyr ago, long after it had formed, differentiated and cooled, all of which happened before 4.45 Gyr ago. The large multi-ringed impact basins on the Moon were formed during that interval. Subsequent volcanic flooding of those basins created the flat, dark 'man in the Moon' maria, whose very flatness testifies to a sharp drop in the cratering rate that ended

the late heavy bombardment (LHB).

Most lunar highland craters were also formed during the cataclysm. Projectiles from the same population should also have struck the nearby Earth; indeed, gravitational scattering of these bodies would ensure that they struck Mercury, Venus and Mars as well. Thus the heavily cratered terrains on Mercury and Mars have been presumed to date from the same LHB, a little less than 4 Gyr ago. (A special population of asteroids called 'vulcanoids' has been hypothesized, though not yet discovered, which could have formed Mercury's cratered terrains more recently.) The LHB must have had profound consequences for the Earth, especially for the ocean and atmosphere. Indeed, until the LHB ended, our planet would have been unsuitable for the origin or survival of life⁵.

The nature of the LHB, and its pervasiveness throughout the Solar System, have been disputed since it was proposed. One reason is the absence of enough relevant planetary samples for dating by analysis of radioactive isotope decay. We also have incomplete knowledge of the present population of comets and asteroids, and even less knowledge about the shorter-lived populations that might have existed in the period following planetary accretion.

A chance break-up of a large comet or asteroid, perhaps by planetary tidal disruption during a close encounter (as happened to Comet Shoemaker–Levy 9 in 1992), could have caused the LHB. The size distribution of LHB craters (dominated by big ones) differs from that produced by present-day asteroids and comets, and indicates such an origin for the projectiles. But several researchers disagree. One view regards the cataclysm itself as mostly an illusion⁶. In this picture there was a fairly gradual decline in the bombardment rate as the remnant planetesimals, which failed to accrete immediately into planets, were swept up. We can't 'see' earlier rock ages or craters because the saturation bombardment masks earlier epochs.

At the opposite extreme⁷ is a theory in which the LHB is so brief (less than 100 Myr) that it must have been caused by projectiles in Earth orbit, perhaps created by the collisional break-up of ancient moons. In that case the LHB was unique to the Earth–Moon system, and the supposition that martian and mercurian craters were formed 4 Gyr ago (on which those planets' geological chronologies are based) is spurious.

So it is tantalizing that argon–argon isotope dating puts the creation of ALH84001 at the same time as the LHB. From a single age, we can hardly deduce a duration for the martian LHB; nor can Mars settle the dispute about whether rock age distributions are biased by saturation. However, a martian LHB could hardly be caused by the hypothetical geocentric projectiles. Argon–argon ages for meteorites⁸ (especially for the basaltic meteorites that are probably derived from the asteroid Vesta, but also for ordinary chondrites) indicate that an LHB also occurred in the asteroid belt during the period 3.5 to 4.1 Gyr ago. The data argue against a short duration for the LHB and probably also against saturation masking. So an LHB cataclysm dominated at least the inner Solar System out to the asteroid belt. Moreover, the signature of the LHB in meteorites implies that it had a very different cause from the usual inter-asteroid collisional break-ups.

The argon–argon radiometric clock can be reset by large, high velocity impacts on bodies greater than several hundred kilometres in diameter⁸. Evidently, there was a real, dramatic increase in the flux of such large projectiles long after the planets had formed. So the LHB was pervasive, and it may have been unrelated to processes of planetary accretion. If a giant comet broke

1. Mittlefehldt, D. A. *Meteoritics* **29**, 214–217 (1994).
2. McSween, H. Y. *Meteoritics* **29**, 757–779 (1994).
3. Ash, R. D., Knott, S. F. & Turner, G. *Nature* **380**, 57–59 (1996).
4. Tera, F., Papanastassiou, D. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **22**, 1–21 (1974).
5. Maher, K. K. & Stevenson, D. J. *Nature* **331**, 612–614 (1988).
6. Hartmann, W. K. *Icarus* **24**, 181–187 (1975).
7. Ryder, G. *Eos* **71**, 313 & 322–323 (1990).
8. Bogard, D. D. *Meteoritics* **30**, 244–268 (1995).

up four billion years ago, maybe another comet could create another LHB in the future.

However, ALH84001 has not yet told its full story. Measurements of carbon isotopes in several SNC meteorites were reported recently at a meeting in Houston. One of several possible interpretations of the unexpected data is that ALH84001 may not be from the same parent body as other

SNCs — that is, not from Mars! Until samples are eventually returned by a mission to Mars, we will have to hope that Antarctic meteorite expeditions will find more SNC meteorites to further elucidate the red planet's early history. □

Clark R. Chapman is at the Southwest Research Institute, 1050 Walnut Street, Boulder, Colorado 80302, USA.

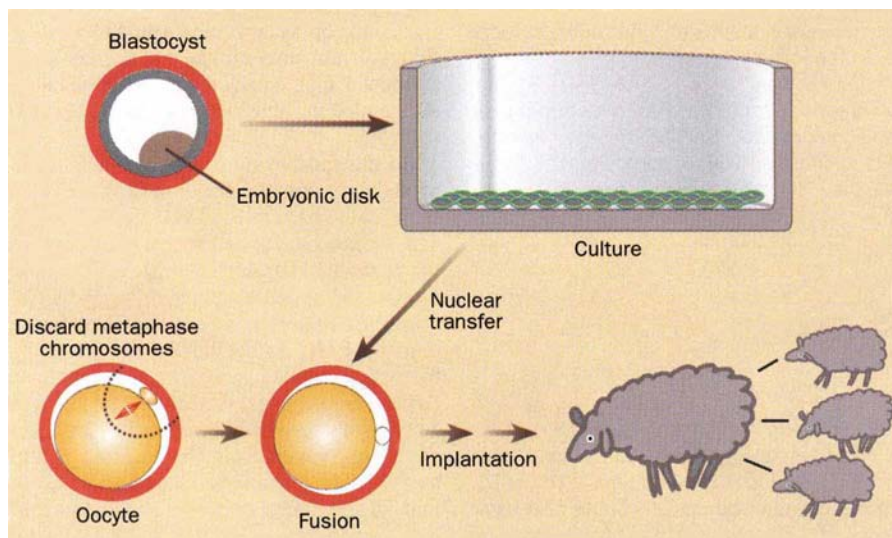
EMBRYOLOGY

Lambing by nuclear transfer

Davor Solter

ALMOST all the exciting experiments that have increased our understanding of the genetic control of mammalian development have been done on the mouse. Attempts to extend this success to other

embryos. This cell line was totipotent, a condition in which cells retain their potential to differentiate into all the different cell types that make up the adult organism. Then, they carefully removed all chromo-



The process by which Campbell and colleagues¹ cloned lambs by nuclear transfer. Embryonic disks were cultured on a feeder layer and permanent cell lines established. The authors then enucleated recipient oocytes by removing a small amount of cytoplasm containing the metaphase plate, and fused them with a single cell from the totipotent cell line. Reconstructed embryos were transferred to temporary recipient ewes and their development was assessed after seven days. Finally, morulae and blastocysts were implanted in other ewes and allowed to develop to term.

mammalian species have generally foundered, so the cloning of identical sheep by nuclear transfer from cultured cells, described on page 64 of this issue by Campbell and co-workers¹, is cause for celebration. Aside from its intrinsic biological interest, this achievement opens up the possibility of manipulating the sheep's genes before cloning them.

The first description of a reliable nuclear transfer technique² was soon followed by reports of its limitations (reviewed in ref. 3). One reason why our understanding of the events inside the cell that follow nuclear transfer has been slow is that mice have proved surprisingly difficult to clone by this method.

In their experiment, Campbell *et al.* first derived a cell line from early sheep

somal material from a recipient oocyte (see figure), fused this oocyte with a totipotent cell that had been through several passages in culture, and grew their manipulated embryos *in vivo*. The transplanted nuclei remained totipotent and embryos that developed to the morula/blastocyst stage were transferred to the uteri of ewes; several lambs were born and grew normally. It is quite fitting that this cloning experiment worked for sheep, because it was in sheep that nuclear transfer of cleavage-stage blastomere nuclei into enucleated eggs first resulted in the birth of normal lambs⁴.

Once permanent cell lines with totipotent nuclei are established, there is in principle no limit to the genetic alterations that can be made: for example, transgenic

strains of sheep and other agricultural animals could readily be established by deleting or modifying endogenous genes. But there are still several problems to be resolved before the technique can realize its full potential.

The overall success rate (live births compared to the number of manipulated embryos) is disappointingly low. The progress *in vivo* to the blastocyst stage and beyond needs to be improved. It is not clear whether the high failure rate is caused by the accumulation of assorted deleterious factors, perhaps caused by repeated manipulations, or whether it reflects heterogeneity within the cell line. The latter is unlikely as the success rate does not change with successive passages, but other, preferably cloned, totipotent cell lines need to be investigated to clarify this point. Knowledge of sheep (and cow and pig) genetics is essential for gene targeting, but at the moment this is rather limited.

Effective nuclear transfer, with development to term of the manipulated egg, must depend on adequate functional reprogramming of the donor nucleus. Macromolecules (messenger RNAs and proteins) stored in oocytes only support mammalian development for a relatively short time (as measured by the number of cell divisions), and the shorter this period, the less time there is for reprogramming. As the embryo undergoes more changes as it grows, the cells from older embryos will require more reprogramming time and will be less likely to reprogramme completely. Compatibility between the recipient cytoplasm and donor nucleus, still poorly understood, plays a part too.

The scant data available indicate that there are at least two factors that contribute to the success of nuclear transfer. First, ovulated oocytes are much better recipients than zygotes, either because there is more time for reprogramming or for some reason their cytoplasm is more suitable. For example, cytoplasmic factors necessary for chromosomal remodelling and genome activation may exist that are titrated out after fertilization, perhaps by binding to replicating DNA or by time-dependent degradation. Second, donor nuclei caught either in the G1 (refs 5, 6) or G0 (ref. 1) phase of the cell cycle result in much better development than those in late S or G2 phase (although there is at least one dissenting report⁷). This intuitively makes sense — it would seem to be easier to reprogramme a genome that is open for and undergoing replication.

Imprinting that results in functional differences in the male and female genome, and the need for both in successful development⁸, could complicate the reprogramming of the genome after nuclear transfer. Setting aside the unlikely possibility of an absence of imprinting in sheep, the achievement of Campbell *et al.* indicates that the cells they used as nuclear donors

retained enough of the functional imprint to support development. This explanation is borne out by the paternal X chromosome of mouse embryonic stem cells, which keeps the imprinting necessary to ensure its preferential inactivation in extraembryonic membranes⁹. On the other hand, a partial loss of imprinting in the cells used by Campbell *et al.* might have contributed to the low rate of development of manipulated embryos.

The benefits of being able to clone valuable farm animals are obvious, although so far the technical and biological problems have limited the potential impact. The results of Campbell and colleagues, the production of calves as a result of nuclear transfer from cultured inner-cell mass cells¹⁰, and the isolation of embryonic stem cell lines from primates¹¹, are all indicators that the technical know-how is to hand to produce clones of different mammals. Cloning mammals from adult cells will be

considerably harder, but can no longer be considered impossible: it might be a good idea to start thinking how we are going to make use of such an option. □

Davor Solter is at the Max Planck Institute for Immunobiology, Stuebeweg 51, 79108 Freiburg, Germany.

1. Campbell, K. H. S., McWhir, J., Ritchie, W. A. & Wilmut, I. *Nature* **380**, 64–66 (1996).
2. McGrath, J. & Solter, D. *Science* **220**, 1300–1302 (1983).
3. Sun, F. Z. & Moor, R. M. *Curr. Topics dev. Biol.* **30**, 147–176 (1995).
4. Willadsen, S. M. *Nature* **320**, 63–65 (1986).
5. Cheong, H.-T., Takahashi, Y. & Kanagawa, H. *Biol. Reprod.* **48**, 958–963 (1993).
6. Otaegui, P. J. *et al. Molec. Repr. Dev.* **39**, 147–152 (1994).
7. Smith, L. C., Wilmut, I. & Hunter, R. H. F. *J. Reprod. Fertil.* **84**, 619–624 (1988).
8. Solter, D. *A. Rev. Genet.* **22**, 127–146 (1988).
9. Tada, T., Tada, M. & Takagi, N. *Development* **119**, 813–821 (1993).
10. Sims, M. & First, N. L. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6143–6147 (1994).
11. Thomson, J. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **92**, 7844–7848 (1995).

DEVELOPMENTAL BIOLOGY

The brain organization

Siew-Lan Ang

THE paper by Crossley *et al.* on page 66 of this issue¹ takes us one step forward in answering a central question in developmental biology — how the regional identity of the vertebrate brain is specified along the anterior/posterior axis. This elegant study shows that *Fgf8*, a member of the fibroblast growth factor family, has a key part in midbrain development.

The first phase in the induction and early regionalization of the neural tube, the structure that develops into the brain and spinal cord, has been extensively studied in the frog, chick and more recently the mouse^{2–4}. It involves inductive interactions during gastrulation between two germ lay-

ers, mesoderm and ectoderm. These early interactions subdivide the front part of the neural tube into broad regions, consisting of the forebrain (prosencephalon), midbrain and anterior hindbrain (mesencephalon–metencephalon; mes–met) and posterior hindbrain (myelencephalon), that are distinguishable by the expression of region-specific genes. In particular, the mes–met region is characterized by the overlapping expression of the homeobox-containing genes *En1* and *Otx2*, the paired-box-containing gene *Pax2*, and the signalling molecules *Wnt1* and *Fgf8* (refs 5–7; *a* in the figure).

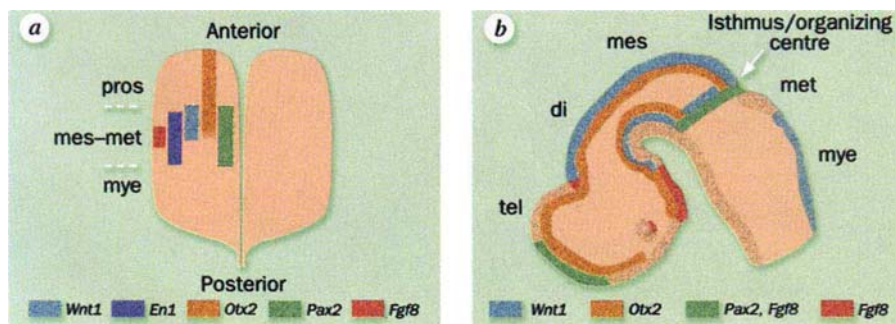
During a second phase of patterning of

the mes–met region, the mesencephalon and metencephalon acquire distinct fates, marked by progressive restriction of *Otx2* and *Wnt1* to the mesencephalic side and of *Pax2* and *Fgf8* to the metencephalic side of a constriction between these two territories called isthmus^{5–7} (*b* in the figure). Mesencephalon and metencephalon will subsequently differentiate into midbrain and cerebellum respectively. Experiments involving the transplantation of embryonic brain tissue in chicks have shown that this second phase of patterning could depend on signals working in the plane of the neural tube^{8,9}. In particular, these studies have revealed a central role for cells of the isthmus. When isthmus tissue is transplanted in the caudal forebrain, the host tissue is transformed into an ectopic midbrain that is the mirror image of the normal midbrain¹⁰. The isthmus tissue thus appears to serve as an organizing centre, producing a signal that can induce and pattern midbrain tissue.

Organizing centres seem to be involved in the development of several structures in both vertebrate and invertebrate species. According to a model proposed by Meinhardt¹¹, the first step in establishing an organizing centre involves the specification of two different populations of cells in adjacent territories. Subsequently, cooperation between these two populations is required to produce a signalling molecule at their common border. This idea has received support from studies of embryonic segmentation¹² and wing and leg imaginal disk development in *Drosophila*^{13,14}.

The demonstration of an organizing centre at the isthmus of the met–mes region, and of adjacent expression of *Wnt1* and *Fgf8* within this region (*b* in the figure), aroused great interest. Crossley *et al.*¹ have now investigated the role of *Fgf8* in development of the mes–met region using grafting experiments in chicks. First, they showed that *Fgf8* is expressed in the isthmus region of chick embryos at the right time to be involved in the development of the organizing centre. They then implanted beads soaked in recombinant FGF8 protein into the caudal forebrain of chick embryos and made the striking observation that FGF8 produced the same effect as transplantation of isthmus tissue — in 17 out of 18 cases, treatment with the beads resulted in the induction of an ectopic midbrain with a mirror-image polarity compared to the normal midbrain. FGF8 also induced ectopic expression of *Wnt1*, *En2* and *Fgf8* itself, three genes that are normally expressed in the mes–met region. Together, these results demonstrate that FGF8 can induce a complete midbrain by establishing an ectopic isthmus-like organizing centre in the forebrain.

From analysis of mouse mutants, it is clear that WNT1 also itself functions in the successive phases of patterning of the mes–met region. For instance, embryos



Expression of the *Wnt1*, *En1*, *Otx2*, *Pax2* and *Fgf8* genes at an early stage (embryo day 8, E8.0) and a late stage (E9.5) of mouse brain patterning. *a*, Dorsal neural plate at E8.0. Overlapping areas of expression of *Wnt1*, *En1*, *Pax2*, *Otx2* and *Fgf8* in the mes–met region are indicated. The domain of *Fgf8* expression shown corresponds to its expression at E8.5 according to Crossley and Martin⁷, as earlier expression of *Fgf8* in the neural plate has not been described. The posterior limit of *Otx2* is fuzzy at this stage, as depicted by fading colour at its caudal end. *b*, Mouse brain at E9.5 in lateral view. *Wnt1* and *Otx2* are expressed in the mesencephalic side, whereas *Fgf8* and *Pax2* are expressed in the metencephalic side of the isthmus. *Pax2* expression in the optic vesicle is not shown. pros, prosencephalon; mes, mesencephalon; met, metencephalon; mye, myelencephalon; tel, telencephalon; di, diencephalon. (*a* and *b* adapted from refs 5 and 7 respectively.)

with a null mutation in *Wnt1* show an early deletion of this region, probably reflecting a loss of expression of *En* genes⁹, which indicates that WNT1 is involved in early growth and regionalization of mes-met. Studies on embryos homozygous for a less severe *Wnt1* point mutation, *swaying*, provide evidence that WNT1 additionally participates in the second phase of patterning of the region (that is, in maintaining the distinction between the mesencephalon and metencephalon)⁶. Patches of *Otx2*-expressing cells are found in the metencephalon of homozygous *swaying* embryos, showing that the mes-met boundary is perturbed. This function in maintaining two differently specified, adjacent territories suggests, according to Meinhardt's model, that *Wnt1* is directly involved in establishing the mes-met organizer.

The induction of a patterning activity by FGF8-bead implants could thus be explained by the ectopic activation of *Wnt1*. The new data¹ constitute strong evidence that endogenous FGF8, which is present in the mesoderm⁷ and neuroectoderm of the mes-met region (*a* in the figure), induces and/or maintains *Wnt1* and *En* expression. It will be interesting to find out whether FGF8 also contributes directly to the patterning activity emanating from the isthmus, and whether WNT1 can regulate *Fgf8* expression in the mes-met region (maintenance of boundaries during development is generally thought to rely on reciprocal interactions between two juxtaposed domains^{11,12}).

FGF8 is also found in other signalling centres in the primitive streak, forebrain and branchial arches⁷. The discovery that it has potent inducing activities in the mes-met region¹ and in the limb¹⁵ makes it tempting to speculate that the molecule participates more generally in regulating patterning in the vertebrate embryo. □

Siew-Lan Ang is in the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/Université Louis Pasteur, BP 163, 67404 Illkirch Cedex, France. siew-lan@igbmc.u-strasbg.fr

1. Crossley, P. H., Martinez, S. & Martin, G. R. *Nature* **380**, 66–68 (1996).
2. Kelly, O. G. & Melton, D. A. *Trends Genet.* **11**, 273–278 (1995).
3. Stern, C. D. *FASEB J.* **8**, 687–691 (1994).
4. Ang, S.-L. & Rossant, J. *Development* **18**, 139–149 (1993).
5. Rowitch, D. H. & McMahon, A. P. *Mech. Dev.* **52**, 3–8 (1995).
6. Bally-Cuif, L., Chorley, B. & Wassef, M. *Mech. Dev.* **53**, 23–34 (1995).
7. Crossley, P. & Martin, G. R. *Development* **121**, 439–451 (1995).
8. Bally-Cuif, L. & Wassef, M. *Curr. Opin. Genet. Dev.* **5**, 450–458 (1995).
9. Joyner, A. L. *Trends Genet.* **12**, 15–20 (1996).
10. Martinez, S., Wassef, M. & Alvarado-Mallart, R. M. *Neuron* **6**, 971–981 (1991).
11. Meinhardt, H. *Dev. Biol.* **96**, 375–385 (1983).
12. Ingham, P. W. & Martinez Arias, A. *Cell* **68**, 221–235 (1992).
13. Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
14. Diaz-Benjumea, F. J., Cohen, B. & Cohen, S. M. *Nature* **372**, 175–179 (1994).
15. Crossley, P. H. et al. *Cell* **84**, 127–136 (1996).

Holography with X-rays

Charles S. Fadley and Patrick M. Len

DIRECT three-dimensional imaging of atomic configurations has, until now, only been possible in a limited way, using electrons to provide clues about near-surface structures. A potentially much more powerful technique using X-ray fluorescence is reported by Tegze and Faigel on page 49 of this issue¹, along with the first X-ray hologram and an atomic image derived from it.

X-ray diffraction is the most widely used technique for determining the bulk atomic structures of solids, even those of large biological molecules. In its most common form, a well collimated and monoenergetic X-ray beam is aimed at a single-crystal specimen, and elastic scattering from atoms in the different atomic planes generates sharp spots of intensity outside the specimen (part *a* in the figure). This method is growing ever more powerful with the new high-brightness synchrotron radiation sources that are now becoming available around the world.

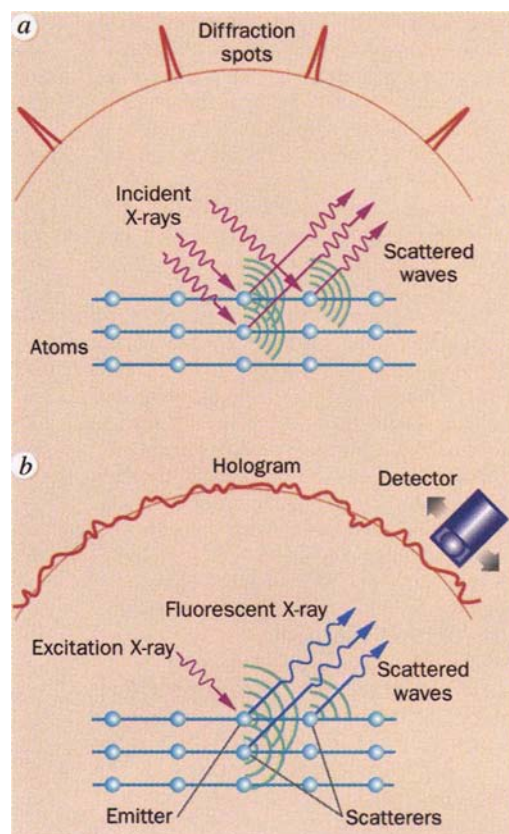
Such structure determinations must still be done by a trial-and-error methodology, in which diffraction spot intensities are calculated for various possible atomic geometries and compared with experimental intensities. The geometry for which theory and experiment best agree is taken as the true structure. This indirect method is made necessary by the 'phase problem' — the phases of individual scattered waves cannot be related to a single reference, but only measured as differences between one another, and this limitation means that the structure cannot be determined uniquely. The difficulty in this kind of search is being certain that all likely structures have been tried so that the actual one is not missed.

Similar problems arise in studies of surface structures using X-ray diffraction, low-energy electron diffraction, or photoelectron diffraction (in which electrons are excited by X-rays from the highly localized core levels centred on each type of atom in the sample). In the latter two cases additional complexities arise because electrons interact much more strongly with atoms than X-rays do, and produce highly anisotropic scattering, phase shifts that are not associated with atomic positions and complex multiple scattering events that are difficult

to analyse theoretically. (The stronger inelastic scattering of electrons does have an advantage for surface studies however, by providing much-increased sensitivity to the first few layers of atoms.)

An idea for getting around these problems, first suggested by Gabor², is to image atoms directly in three dimensions by holography. In holography, the phase problem is solved by letting the scattered waves interfere with a single reference wave. Holographic images using visible light are now a part of daily life, but their resolution is limited by the wavelength of light (several thousand atomic diameters).

A method of going further, so as to achieve atom-scale resolutions, was proposed about a decade ago³. Any localized atomic source of outgoing waves with a well-defined, short wavelength (energetic X-rays or electrons for example) can be thought of as producing a hologram



a, In X-ray diffraction, an incident wave is scattered elastically from many atoms in planes a fixed distance apart to produce sharp diffraction spots where the scattered waves interfere constructively. *b*, In local-source X-ray holography, the formation of a core-electron-level vacancy in an atom by absorption of an X-ray is followed by fluorescent X-ray emission at a longer wavelength. This X-ray in turn scatters elastically off nearby atoms to produce a diffraction pattern or hologram through interference between the fluorescent reference wave and scattered wave components.

through its scattering from near-neighbour atoms, with the unscattered part being used as the reference wave (part *b* in the figure). Suitable sources are core photoelectrons, X-rays emitted in core-level fluorescence processes, and γ -rays emitted by atomic nuclei. This idea was initially expanded on for the case of photoelectrons^{4,5}, and it has now been applied in several forms to surface studies⁶⁻¹⁴.

Electron holographic methods have already become useful for minimizing trial-and-error searches in some aspects of surface structure determinations. But what of holography with X-rays, which are expected to scatter more ideally and thus provide more accurate atomic images? The potential for local-source X-ray holography has already been discussed from a theoretical point of view^{15,16}, but it is only with the work of Tegze and Faigel¹ that the method has been given the first exciting experimental realization.

The authors exposed a single crystal of the perovskite SrTiO_3 to X-rays from a standard source at about 17.4 keV to generate fluorescent X-rays from the strontium atoms in the crystal at about 14.1 keV (corresponding to an atomic-scale wavelength of 0.88 Å). Fluorescence X-rays are sufficiently monoenergetic to produce a holographic interference pattern. The hologram was recorded by rotating the crystal and the detector to measure the intensity profile at 14.1 keV over a large solid angle above the surface (part *b* in the figure).

The strontium atoms in this crystal all occupy the same lattice position, so the outgoing X-rays probe geometrically identical local atomic neighbourhoods and the final hologram is thus a sum of identical holograms for all the strontium atoms in the lattice. Although there is some scattering from titanium and oxygen atoms, scattering from strontium is much stronger because of its higher atomic number (38), so the holographic image should show, to first order, a cubic array of strontium atoms.

One negative aspect of using X-ray scattering is that the diffraction effects are weak, being only 0.1% of the total intensity or less (compared with perhaps 50% for electrons). In spite of this difficulty, Tegze and Faigel were able to measure approximately 0.3% X-ray diffraction effects. Measured intensities were inverted mathematically to form an atomic image, using a Fourier transform-like procedure from optics known as the Helmholtz-Kirchoff formula⁴. The first local-source X-ray holographic images resulting from this procedure are shown on page 51 of this issue.

By choosing to excite a certain ray fluorescence preferentially (for example with a tunable synchrotron radiation X-ray

source) it should be possible to probe the atomic neighbourhood of each type of atom in a multi-component material or molecule. Because synchrotron sources are much brighter than ordinary X-ray tubes, and are getting brighter all the time, it will be possible to reduce the counting times needed to measure weak diffraction patterns accurately and to study atoms that are present in much lower concentrations. With sufficiently fast next-generation detectors one might think of studying the local environment around dopant atoms in semiconductors or metal atoms at active sites in biological molecules.

Localized sources of γ -rays from nuclear decay would produce even stronger scattering and diffraction effects involving the nucleus, and would be sensitive to magnetic or chemical effects via small shifts in γ -ray energy from emitter to scatterer.

One difficulty with the method is the presence of strong Bragg diffraction effects known as Kossel lines, which can appear at certain directions even in the fluorescence X-ray emission pattern, but these effects are expected to have a much sharper angular profile than the local-neighbourhood hologram, and so could be simply filtered out¹⁵. A second problem is the presence of twin images, which might confuse the interpretation through overlaps and destructive interference^{15,16}, but this could be solved if some way is found of obtaining local-source X-ray holograms at several energies^{5,16}. □

Charles S. Fadley and Patrick M. Len are in the Department of Physics, University of California at Davis, Davis, California 95616, and the Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

1. Tegze, M. & Faigel, G. *Nature* **380**, 49-51 (1996).
2. Gabor, D. *Nature* **161**, 777-778 (1948).
3. Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Applications* (eds Attwood, D. T. & Bokor, J.) 361-467 (American Institute of Physics, New York, 1986).
4. Barton, J. J. *Phys. Rev. Lett.* **61**, 1356-1359 (1988).
5. Barton, J. J. *Phys. Rev. Lett.* **67**, 3106-3109 (1991).
6. Saldin, D. K. & de Andres, P. L. *Phys. Rev. Lett.* **64**, 1270-1273 (1990).
7. Harp, G. R., Saldin, D. K. & Tonner, B. P. *Phys. Rev. Lett.* **65**, 1012-1015 (1990).
8. Thevuthasan, S. et al. *Phys. Rev. Lett.* **67**, 469-472 (1991).
9. Thevuthasan, S. et al. *Phys. Rev. Lett.* **70**, 595-598 (1993).
10. Tong, S. Y., Li, H. & Huang, H. *Phys. Rev. Lett.* **67**, 3102-3105 (1991).
11. Tong, S. Y., Li, H. & Huang, H. *Phys. Rev.* **B51**, 1850-1854 (1995).
12. Wu, H. & Lapeyre, G. J. *Phys. Rev.* **B51**, 14549-14553 (1995).
13. Terminello, L. J., Barton, J. J. & Lapianto-Smith, D. A. *J. Vacuum Sci. Technol.* **B10**, 2088-2091 (1992).
14. Terminello, L. J., Barton, J. J. & Lapianto-Smith, D. A. *Phys. Rev. Lett.* **70**, 599-602 (1993).
15. Tegze, M. & Faigel, G. *Europhys. Lett.* **16**, 41-46 (1991).
16. Len, P. M., Thevuthasan, S., Fadley, C. S., Kaduwela, A. P. & Van Hove, M. A. *Phys. Rev.* **B50**, 11275-11278 (1994).

Simple tastes

SPICES were once the most craved of products. For centuries they were the main, highly lucrative, export from Asia to Europe. This is hard to understand these days. Possibly the basic European diet was so bland and boring that almost any price was worth paying to liven it up; more likely, it was always on the edge of putrefaction, and disguise was desperately needed. This strategy, says Daedalus, is chemically simplistic. What is needed is not more additions, but judicious subtraction. He recalls that a poor red wine can be strikingly improved by filtration through activated charcoal, which removes the tannins and coarser flavours. Probably the finest cooking owes its delicacy not to what is present, but what is absent.

So DREADCO biochemists are seeking to identify the key flavour components of bad cooking. They are heroically patronizing 'greasy spoon' restaurants, ancient school refectories and the staff canteens of failing companies, while DREADCO's cooks do their own worst with the oldest and nastiest ingredients. With luck, these essays in culinary disaster will reveal a common chemical motif. A few crucial trace components, perhaps specific smelly amines and thiols from oxidation or bacterial attack, will dominate most of the worst flavours.

The team will then devise a chemical antidote, to sequester and inactivate these compounds. Subtle and delicate chelating agents will probably not work. To withstand the rigours of cooking, an absorbent as stable and robust as activated charcoal will be needed. DREADCO chemists are modifying selected charcoals by incorporating molecular sieve minerals and adding selectively reactive side chains.

DREADCO's 'Antispice' will resemble a fine black pepper. Added to the ingredients, it will raise the tone and appeal of any meal. It may not lift it to the heights of the best cordon-bleu artistry; but it will certainly rescue it from mediocrity, and even from the depths of disaster. It will be welcome in the best kitchens, and even more welcome in the worst ones.

But Daedalus goes further. These days many consumers are very fearful of nasty chemical additives in their diet. So for dedicated diet-worriers, the DREADCO team is seeking a range of 'subtractives' to sequester and neutralize artificial colouring, anti-oxidants, monosodium glutamate, cholesterol, salt and so on. Sadly, a subtractive to absorb all fats would be a major component of the meal in itself; and even Daedalus cannot dream up a simple way of selectively absorbing all calories. David Jones

Transparent light-emitting devices

SIR — Organic light-emitting devices (OLEDs), which emit in the red¹, green² and blue³ spectral regions, have recently become promising candidates for colour flat-panel display pixels owing to their high luminosity and low operating voltage. Furthermore, long device lifetimes⁴ (exceeding 1,000 hours at a video brightness of $\sim 100 \text{ cd m}^{-2}$) have been demonstrated. One unique property of vacuum-deposited OLEDs is that the luminescence band is substantially redshifted from the absorption band⁵, rendering the organic layers highly transparent to their own luminescence and throughout most of the visible spectrum.

Here we demonstrate a new class of OLEDs, which are greater than 70% transparent when turned off, and emit light from both top and bottom surfaces with up to 0.75% quantum efficiency when turned on. Such transparency offers the potential for very high-definition, full-colour displays in which the red (R), green (G) and blue (B) emission layers are placed in a vertically stacked geometry, providing a simple, low-temperature fabrication process, as well as minimum R-G-B pixel size and maximum fill factor. Further applications

of this device include low-voltage, semi-transparent displays for helmet-mounted, windscreen-mounted or other 'head-up' applications.

The transparent OLED (TOLED) structure is shown in Fig. 1a. The device is grown on a glass substrate pre-coated with a thin film of transparent indium tin oxide, with a sheet resistivity of 20Ω per square. Before deposition of the organic films, the substrates were pre-cleaned as discussed elsewhere⁵. Deposition was performed by sublimation, in a vacuum of $<10^{-6}$ torr, of a 200-Å-thick layer of the hole-conducting compound *N,N'*-diphenyl-*N,N'*-bis(3-methylphenyl)-1,1'-biphenyl-4,4'-diamine (TPD), followed by a 400-Å-thick layer of the electron-conducting and highly electroluminescent tris(8-hydroxyquinoline) aluminium (Alq_3).

Figure 1a shows how electron-injecting contact to the device was made by deposition through a shadow mask of a thin layer (50–400 Å) of Mg–Ag alloy (in an approximate atomic ratio of 40 Mg:1 Ag). Finally, the device was capped by a second 400-Å-thick layer of indium tin oxide, sputter-deposited onto the Mg–Ag surface to provide a continuous, transparent conducting surface. The sheet resistance of this layer is 400Ω per square, which acts in parallel to the Mg–Ag sheet resistance and is adequate to serve as an injecting contact for a display pixel.

Figure 1 shows a photograph of a TOLED array with a 100-Å-thick Mg–Ag layer. In Fig. 1b, one TOLED array element is switched on (arrow) and the inactive devices are only faintly visible. They reduce the light from the background by only $\sim 3 \text{ dB}$, even though these images are taken under the condition that the light illuminating the background must pass twice through the TOLED. In Fig. 1c, all the elements are switched off. The transparency as a function of a TOLED from this array is shown in detail in Fig. 2. The device becomes non-transparent at short wavelengths owing to a combination of Mg–Ag absorption and the strong molecular transitions to the $^1\text{L}_a$ and $^1\text{B}_u$ states⁶ of Alq_3 , and at long wavelengths due to absorption by the Mg–Ag. However, the device is 63% transparent at the peak (530 nm) emission wavelength of Alq_3 , and this transparency extends across the visible spectrum.

The inset to Fig. 2 shows the transmission of the Mg–Ag contact at a wavelength of 530 nm, with Mg–Ag film thicknesses ranging from $(50 \pm 10) \text{ Å}$ to $(400 \pm 10) \text{ Å}$. The thinnest Mg–Ag layer used thus far in a working device is 75 Å thick, corresponding to a contact trans-

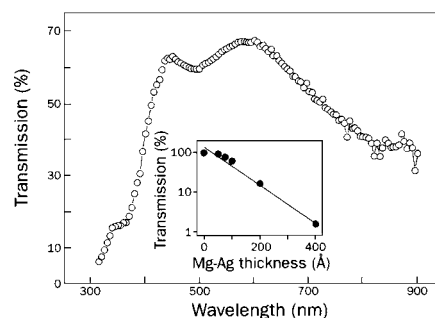


FIG. 2 Transmission spectrum of a TOLED with a 100-Å-thick Mg–Ag electrode as a function of wavelength. Inset, Transmission as a function of Mg–Ag contact thickness at a wavelength of 530 nm. The slope of the straight line fit to the data gives an optical absorption coefficient of the Mg–Ag of $\alpha = 1.1 \times 10^6 \text{ cm}^{-1}$, consistent with a calculated skin depth of 125 Å.

parency of 81% and a total device transparency of 71%. The real and imaginary coefficients of the refractive index of Mg at 530 nm are $n=0.57$ and $k=3.47$, respectively⁷. Neglecting the small amount of Ag in the electrode, we calculate that an 81% transmissive film of Mg should be 15 Å thick. This suggests that approximately $(60 \pm 10) \text{ Å}$ of the Mg–Ag electrode is compositionally changed in the deposition process, presumably due to a redox reaction with indium tin oxide during sputtering. This explains the higher measured transparency, the departure of our transmission versus thickness data from a straight line, and our inability to produce working devices with less than a 75-Å-thick layer of Mg–Ag.

In conclusion, we have demonstrated the first transparent, thin-film organic light-emitting devices. This is the crucial first step towards realizing high-definition, full-colour and head-up displays using organic materials.

V. Bulović

G. Gu

P. E. Burrows

S. R. Forrest

Advanced Technology Center for Photonics
and Optoelectronic Materials,
Department of Electrical Engineering,
Princeton University, Princeton,
New Jersey 08544, USA

M. E. Thompson

Department of Chemistry,
University of Southern California,
Los Angeles, California 90089, USA

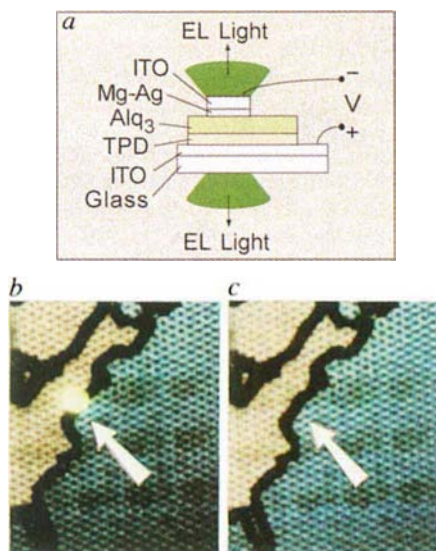


FIG. 1 a, Schematic diagram of the TOLED structure (ITO, indium tin oxide). b, An array of TOLEDs with one device turned on (green spot) against a backdrop of the eastern US coastline. Typical operating conditions for the 1-mm-diameter devices are 10^{-4} A and 10 V drive voltage. The emission spectra from both surfaces are similar to those of conventional Alq_3 -based devices^{2,5}, after accounting for absorption in the Mg–Ag layer. The total quantum efficiency of light emission from this device is 0.75%. Approximately 10% higher intensity is emitted from the substrate than from the top contact surface. c, The same array with all devices turned off, demonstrating the device transparency.

1. Takada, N., Tsutsui, T. & Saito, S. *J. appl. Phys.* **33**, L863–L866 (1994).
2. Tang, C. W. & VanSlyke, S. A. *Appl. Phys. Lett.* **51**, 913–915 (1987).
3. Adachi, C., Tsutsui, T. & Saito, S. *Appl. Phys. Lett.* **56**, 799–801 (1990).
4. Burrows, P. E. et al. *Appl. Phys. Lett.* **65**, 2922–2924 (1994).
5. Burrows, P. E. & Forrest, S. R. *Appl. Phys. Lett.* **64**, 2285–2287 (1993).
6. Garbuzov, D. Z., Bulović, V., Burrows, P. E. & Forrest, S. R. *Chem. Phys. Lett.* (in the press).
7. O'Bryan, H. M. *J. opt. Soc. Am.* **26**, 122–127 (1936).

HIV reverse transcription in yeast

SIR — The study of human immunodeficiency virus (HIV) reverse transcriptase/ribonuclease H (RT/RH) and other medically important reverse transcriptases would be greatly enhanced by simple, safe and inexpensive *in vivo* assays. We combined the reverse transcription indicator gene *his3AI* with hybrid Ty/HIV elements to demonstrate that HIV-1 RT/RH can substitute for the RT/RH of the Ty1 retrotransposon in *Saccharomyces cerevisiae*.

S. cerevisiae strains harbour retrotransposons (Ty elements) which replicate like retroviruses¹. Ty encodes an RT/RH necessary for replicative transposition and an integrase required for insertion at novel sites. The plasmid pGlyhis3AI (*a* in the figure) was developed for analysis of retrotransposition². Splicing and reverse

transcription of the galactose-inducible Tyhis3AI messenger RNA yields Ty complementary DNA carrying a functional *HIS3* gene that can be integrated into the genome. Ty cDNA can also undergo homologous recombination with endogenous Ty elements. Both transposition and cDNA recombination result in His⁺ cells.

Hybrid Ty/HIV-1 elements (HARTs; *b* in the figure) contain HIV-1 RT/RH adjacent to the Ty protease cleavage site separating Ty1 integrase and RT/RH, completely replacing Ty RT/RH. Yeast expressing these elements produce HIV-1 RT/RH as a protein of relative molecular mass 66,000, suggesting that processing is efficient (D.V. N. and S. Moore, unpublished results). The *in vivo* activity of the HARTs is readily detected in strains defective in endogenous Ty expression (*spt3*; ref. 3) by the production of His⁺ cells (see table). Frameshift mutation of the HIV-1 open reading frame, or point mutations that disrupt either the polymerase or RNase H activities of HIV-1 RT/RH (ref. 4), decrease the production of His⁺ cells by 100–500-fold. Residual activity in the RT/RH mutants probably reflects low-level expression of endogenous Ty elements and can be further reduced by performing the assay at 30 °C, when Ty transposition, but not the activity of the Ty/HIV-1 hybrids, is inhibited⁵.

Ty transposition occurs by integrase activity, is independent of homologous recombination and, indeed, occurs in cells (*rad52*) incapable of homologous recombination^{6,7}. In contrast, the HART elements are dependent on *RAD52* for production of His⁺ cells (see table), suggesting that the cDNA produced by HIV-1 RT/RH is not a substrate for Ty integrase. It is, in fact, unlikely that HIV-1 RT/RH recognizes the Ty signals (polypurine tract and transfer RNA primer binding site) necessary to generate a complete element with correct ends for transpositional integration. We propose that HART cDNAs recombine into the genome by homology with endogenous Ty elements, as has been demonstrated for integrase-defective Ty elements and Ty elements that cannot produce correctly processed ends⁷.

Although an *Escherichia coli*-based *in vivo* assay for the DNA-dependent DNA polymerase activity of HIV-1 RT/RH has been described⁸, our HARTs provide the first efficient *in vivo* assays for RNA-

dependent polymerase and RNase H activities of HIV-1 RT/RH outside the intact virus. Screening for drugs that inhibit the RT/RH of HIV-1 or other retroviruses with a simple genetic assay in yeast is therefore possible. This assay may also be useful in characterizing drug-resistant variants of HIV-1 RT/RH.

D. V. Nissley

D. J. Garfinkel

J. N. Strathern

Gene Regulation and Chromosome Biology Laboratory,
NCI-Frederick Cancer Research and Development Center,
ABL-Basic Research Program,
PO Box B, Frederick,
Maryland 21702, USA

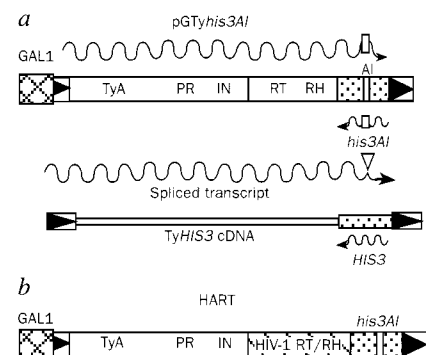
1. Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. *Cell* **40**, 491–500 (1985).
2. Curcio, M. J. & Garfinkel, D. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 936–940 (1991).
3. Winston, F., Durbin, K. J. & Fink, G. R. *Cell* **39**, 675–682 (1984).
4. Boyer, P. L., Ferris, A. L. & Hughes, S. H. *J. Virol.* **66**, 1031–1039 (1992).
5. Paquin, C. E. & Williamson, V. M. *Science* **226**, 53–55 (1984).
6. Melamed, C., Nevo, Y. & Kupiec, M. *Molec. cell. Biol.* **12**, 1613–1620 (1992).
7. Sharon, G., Burkett, T. J. & Garfinkel, D. J. *Molec. cell. Biol.* **14**, 6540–6551 (1994).
8. Kim, B. & Loeb, L. A. *Proc. natn. Acad. Sci.* **92**, 684–688 (1995).

Why are toads right-handed?

SIR — We were pleased to see the report by Bisazza *et al.*¹ indicating that toads are preferentially right-handed. Much of Bisazza *et al.*'s data came from tests where toads preferentially used their right hand to wipe away material stuck to their head and face. We offer here a simple adaptive explanation for why these animals preferentially used that hand for this grooming activity.

Anurans have strong emetic reflexes². Toxic material in the stomach provokes vomiting, and when anurans vomit, they not only regurgitate stomach contents but also the stomach itself! Because the anuran stomach is asymmetric — with a shorter mesentery on the right than the left side — the stomach is tethered to that side. Consequently, the prolapsed stomach always hangs out the right side of the mouth³.

We have watched and videotaped emesis in anurans from many different families and genera², including *Bufo*. In a stereotypic fashion, frogs and toads wipe away remaining vomitus from the surface of the prolapsed stomach before re-swallowing it³. Any anuran that does



The *his3AI* indicator gene and hybrid Ty1/HIV-1 RT/RH elements. *a*, pGlyhis3AI, the inducible galactose-promoter-driven Ty1 element marked with *his3AI* (ref. 2), serves as an indicator of passage through an RNA intermediate and reverse transcription. Ty RNA (wavy line), carrying an antisense copy of *his3AI* containing a spliceable artificial intron (AI; open rectangle) interrupting *his3*, is first spliced and then reverse-transcribed, generating TyHIS3 cDNA. Integration or homologous recombination⁷ of this cDNA results in histidine prototrophy. *b*, The RT/RH domain of HIV-1 was used to replace Ty1 RT/RH, resulting in HART elements.

HIV-1 RT/RH HYBRID ELEMENTS ARE ACTIVE IN YEAST				
		Frequency of His ⁺		
		20°C	20°C	30°C
HIV RT/RH	<i>spt3</i>	<i>RAD52</i>	<i>spt3 rad52</i>	<i>spt3 RAD52</i>
Ty	—	2 × 10 ⁻²	7 × 10 ⁻²	
HART 1	Wild type	6 × 10 ⁻³	7 × 10 ⁻⁵	
HART 1-m	Frameshift	1 × 10 ⁻⁵	8 × 10 ⁻⁶	
HART 21	Wild type	1.8 × 10 ⁻³		2.7 × 10 ⁻³
HART 22	pol ⁻	6.7 × 10 ⁻⁶		<3.0 × 10 ⁻⁷
HART 23	rh ⁻	1.6 × 10 ⁻⁵		<1.5 × 10 ⁻⁷

HIV-1 RT/RH-mediated reverse transcription was monitored by selection of histidine prototrophy following the induction of HART expression as described². The frequency of HART-mediated histidine prototrophy were determined in *RAD52* (DG1251: MATα *ura3-167 trp1-GB spt3-101 his3Δ200*) and *rad52* (DG1286: MATα *ura3-167 trp1-GB spt3-101 his3Δ200 rad52-GB*) strains at 20 °C and 30 °C.

1. Bisazza, A., Cantalupo, C., Robins, A., Rogers, L. J. & Vallortigara, G. *Nature* **379**, 408 (1996).
2. Naitoh, T., Imamura, M. & Wassersug, R. *J. comp. Biochem. Physiol.* **100C**, 353–359 (1991).
3. Naitoh, T., Wassersug, R. J. & Leslie, R. A. *Physiol. Zool.* **62**, 819–843 (1989).

not do this risks re-swallowing noxious material.

Because the stomach is always on the right, it should be no surprise that anurans always use their right hand for this 'gastric grooming'. In fact, the prolapsed stomach normally lies posterior, near the angle of the jaw, and beyond the reach of the left hand.

None of our observations detracts from the conclusions of Bisazza *et al.*, but they do suggest that in a search for the evol-

utionary origin of brain lateralization in higher vertebrates, one should take into consideration natural asymmetries in viscera and visceral functions.

Tomio Naitoh

Department of Biology,
Shimane University,
Matsue 690, Japan

Richard Wassersug

Department of Anatomy and Neurobiology,
Dalhousie University,
Halifax, Nova Scotia B3H 4H7, Canada

tolerant. Transgenic, BASTA-tolerant *B. campestris*-like plants with 20 chromosomes and a high pollen fertility were also found at the experimental site the following spring, among plants that had germinated from seeds shed before harvest the previous year (table).

We have shown that *B. campestris*-like plants with 20 chromosomes, a high pollen fertility, and carrying a transgene from oilseed rape, can be produced as early as in the first-backcross generation, by interspecific backcrossing under field conditions. The occurrence of fertile, transgenic weed-like plants after just two generations of hybridization and backcrossing suggests a possible rapid spread of genes from oilseed rape to the weedy relative *B. campestris*, and this should be taken into account when considering the consequences of transferring new traits to oilseed rape.

Thomas R. Mikkelsen

Bente Andersen

Rikke Bagger Jørgensen

Plant Genetics Section,
Environmental Science and Technology
Department,
Risø National Laboratory,
DK-4000 Roskilde, Denmark

1. Raybould, A. F. & Gray, A. J. *J. appl. Ecol.* **30**, 199–219 (1993).
2. Keeler, K. H. & Turner, C. E. In *Risk Assessment in Genetic Engineering* (eds Levin, M. A. & Strauss, H. S.) 189–218 (McGraw-Hill, New York, 1991).
3. Linder, C. R. & Schmitt, J. *Molec. Ecol.* **3**, 23–30 (1994).
4. Jørgensen, R. B. & Andersen, B. *Am. J. Bot.* **81**, 1620–1626 (1994).
5. Jørgensen, R. B., Andersen, B., Landbo, L. & Mikkelsen, T. R. *Acta hort.* (in the press).
6. Stace, A. C. *New Flora of the British Isles* (Cambridge Univ. Press, 1991).

Soil classification

SIR — J. H. Shorter *et al.* (*Nature* **377**, 717–719; 1995) deserve congratulations for showing by field and laboratory experiments that methyl bromide in some surface soils is subject to rapid bacterial degradation, and that the residence time of this chemical is presumably much shorter than previously estimated. Hence, the continuous use of methyl bromide as a soil fumigant may be much less dangerous for ozone depletion than hitherto assumed.

As Shorter *et al.* point out, more experiments are required before their significant result can be extrapolated generally. I would like to raise a different point arising from their paper, that of the characterization of the soils tested, which are described as "temperate forest soil type", "temperate agricultural surface soil", "sandy boreal forest surface soil" and "tropical forest surface soil", without additional details about their pH, nature of clay minerals, physical properties or biological activity.

Soils are by nature extremely variable, even over small distances; it would be

The risk of crop transgene spread

SIR — An environmental concern regarding the cultivation of genetically modified crop plants is whether transgenes can be transferred to wild or weedy relatives by hybridization and backcrossing, potentially resulting in plants with enhanced invasiveness or weediness^{1,2}. The problem has been addressed for several crop-weed systems¹, but attention has focused on the initial crop-weed hybridization, and there is a lack of experimental data concerning the subsequent stages in the gene transfer process³.

We have studied the introgression of genes from oilseed rape (*Brassica napus*) to the weedy relative *Brassica campestris* (*B. rapa*). The two species hybridize spontaneously, and hybrids have been found in natural populations^{4,5}. Here we report that transgenic, herbicide-tolerant weed-like plants with *B. campestris* morphology, the same number of chromosomes as *B. campestris*, and high fertility, are produced as early as the first-backcross generation in field experiments where transgenic, herbicide-tolerant, interspecific hybrids are grown together with *B. campestris*.

We investigated hybridization and backcrossing between transgenic, BASTA (glufosinate)-tolerant oilseed rape (genome AACC; $2n=38$ chromosomes)

and *B. campestris* (genome AA, $2n=20$ chromosomes) in field experiments. Transgenic interspecific hybrids were produced spontaneously on both parental species, confirming previous observations^{4,5}. For backcrossing, we grew transgenic interspecific hybrids in small plots together with *B. campestris*, and harvested plants at maturity. Contrary to previous observations⁶, interspecific hybrids were not sterile, but produced an average of more than 450 seeds per plant.

We germinated more than 4,000 seeds harvested from 32 interspecific hybrids, and grew 865 randomly chosen transgenic, BASTA-tolerant plants to flowering. Of these plants, we analysed 44 displaying a *B. campestris*-like morphology. We found *B. campestris*-like plants with 20 chromosomes and a pollen fertility of greater than 90%, both characteristics of wild *B. campestris*, among the selected plants (see table). We crossed four of the plants possessing *B. campestris* characteristics with genuine *B. campestris* individuals. The resulting seeds had a pronounced seed dormancy (as weedy *B. campestris*), and the crosses resulted in an average of 6.4 viable second-backcross (BC_2) offspring plants per pollination (table). Of the 416 BC_2 plants obtained, 42% were BASTA-

FIRST-BACKCROSS PLANTS DISPLAYING *B. CAMPESTRIS* CHARACTERISTICS

Plant	No. of chromosomes	Pollen fertility	BC_2 production (no. of pollinations)
Plants grown from harvested seeds:			
B131	20 (20–21)	—	—
AA22	20 (20–21)	94%	7.6 (16)
AA23	21	96%	12.7 (7)
FF13	20 (20–21)	97%	10.1 (18)
HH19	21 (20–21)	95%	1.0 (24)
Plants collected at experimental site:			
P4-52	20 (19–21)	95%	—
P4-57	20 (20–22)	95%	—
<i>B. campestris</i> (parental species)	20	>90%	

Pollen fertility was estimated by staining with cotton-blue, and chromosome counts were performed on root tips. The two plants collected at the experimental site were found among 11 plants selected on their *B. campestris* morphology from BASTA-tolerant plants growing at the experimental site. BC_2 production is the average number of viable offspring plants produced per pollination when pollinated with genuine *B. campestris*.

impossible to resample or retest them on the basis of the information provided by Shorter *et al.* in their paper. Soil classification facilitates their study, use and understanding by giving soil types specific names. This represents a shorthand for numerous characteristic properties. Researched soil bodies should be properly identified and classified, in a similar fashion to vegetation, biomes, minerals or rocks, to facilitate recognition and data transfer.

Each of the biomes listed in Table 2 of the paper by Shorter *et al.* includes many different soil types, not necessarily possessing the same properties and reactivity as those tested. The degradation process is, after all, a function of the soil environment; hence, an appropriate soil identification is desirable, not only when sampling the soils, but also when extrapolating such data to other soil landscapes and regions.

Soil scientists recognize several hundred main soil types, and commonly use two or three worldwide soil classification schemes, mostly the Soil Map of the World scheme from the Food and Agriculture Organization (FAO), or the US Soil Taxonomy. I have recently updated soil classi-

fications (*Catena* 24, 233–241; 1995), listing and reviewing the newest publications. The areal extent of the major global soil groups has been published in various places and can be obtained from computerized databases. The most recent figures on the areal soil distribution for regional or global extrapolation of the major mapped soil units can be obtained from the Land and Water Development Division of FAO, Rome, with references for the FAO soil classification scheme, and from the Division of Soil Survey, Natural Resources Conservation Service, US Department of Agriculture, Washington DC, for the soil taxonomy scheme.

Leading soil-science journals require that soils reported in publications be classified according to one of the two or three internationally recognized and used soil-taxonomy schemes. My appeal is thus that those submitting data on soils to *Nature* and similar multidisciplinary journals follow this sound rule to enhance international communication.

Dan H. Yaalon

*Institute of Earth Sciences,
The Hebrew University,
Givat Ram Campus,
Jerusalem 91904, Israel*

A covalently crosslinked histone

SIR — Post-translational modification of core histones is essential to processes requiring chromatin remodelling. We report here a novel histone modification, involving crosslinking between a glutamine residue of histone H2B (ref. 1) and a lysine residue of histone H4 (ref. 2), in sperm of the starfish, *Asterina pectinifera*.

A unique nuclear protein, designated p28, was isolated from a histone fraction prepared from starfish testes. Amino-acid analysis of the acid-hydrolysed p28 revealed that the amino-acid profile was nearly the same as the sum of the amino acids in testicular histones H2B and H4 (data not shown). Furthermore, electro-

spray ionization mass spectra showed the relative molecular mass of p28 to be 24,654, which is 17 mass units less than the sum of the observed relative molecular masses of histones H2B (13,362) and H4 (11,309).

The structure of p28 was investigated by digestion with *Achromobacter lyticus* protease I or *Staphylococcus aureus* V8 protease. Analysis of the proteolytic fragments after separation by high-performance liquid chromatography suggested that they might be identical to the amino-acid sequences of both H2B and H4, except for the amino-terminal portion (solid box in *a* in the figure): K8, derived from *A. lyticus* protease I digestion. Amino-acid analysis of K8 using constant boiling HCl indicated a composition of Ser:Glu:Gly:Lys:Arg of 1:1:5:3:1, which corresponds to Gly 8–Lys 10 of H2B and Ser 1–Lys 8 of H4. Sequence analysis of K8 yielded Gly-X-Lys, where X denotes a missing residue. Sequence analysis after treatment with trifluoroacetic acid³ gave the sequence Ser-Gly-Arg-Gly-X-Gly-Gly-Lys, corresponding to the N-terminal sequence of H4, in addition to Gly-X-Lys of H2B.

The positive-ion fast-atom bombardment tandem mass spectrum of K8 not only gave the product ions arising from the loss of the constituent amino-acid residues, but also those involved in the crosslink, ϵ -(γ -glutamyl)lysine (see figure). Finally, K8 was digested sequentially with subtilisin, leucine aminopeptidase and carboxypeptidase Y. Amino-acid analysis of the digest revealed the presence of one molecule of ϵ -(γ -glutamyl)lysine.

From these data, the structure of p28 was unambiguously shown to be a heterodimer composed of histones H2B and H4, probably formed through an acyl transfer reaction catalysed by transglutaminase⁴ between the Gln 9 of H2B and the Lys 5 of H4. We believe this to be the first demonstration of a covalently crosslinked histone heterodimer in eukaryotes.

Takahiko Shimizu

Keiko Hozumi

Shintaro Horiike

Kazuo Nunomura

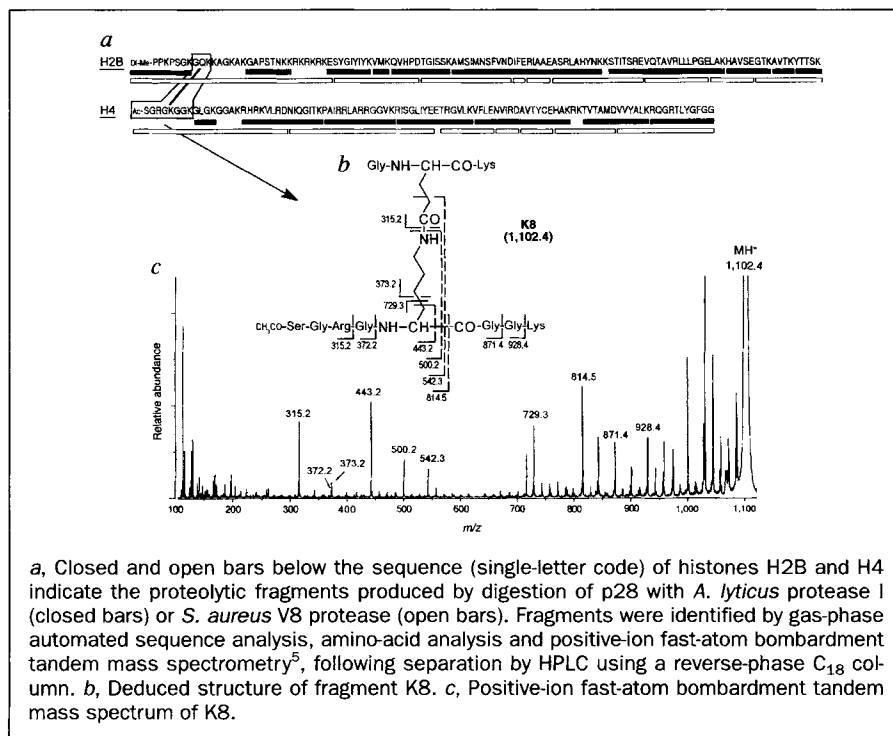
Susumu Ikegami

*Department of Applied Biochemistry,
Hiroshima University, Kagamiyama 1-4-4,
Higashi-hiroshima, Hiroshima 739, Japan*

Toshifumi Takao

Yasutsugu Shimonishi

*Institute for Protein Research,
Osaka University, Osaka 565, Japan*



a, Closed and open bars below the sequence (single-letter code) of histones H2B and H4 indicate the proteolytic fragments produced by digestion of p28 with *A. lyticus* protease I (closed bars) or *S. aureus* V8 protease (open bars). Fragments were identified by gas-phase automated sequence analysis, amino-acid analysis and positive-ion fast-atom bombardment tandem mass spectrometry⁵, following separation by HPLC using a reverse-phase C₁₈ column. **b**, Deduced structure of fragment K8. **c**, Positive-ion fast-atom bombardment tandem mass spectrum of K8.

1. Martinage, A., Briand, G., Van Dorsselaer, A., Turner, C. H. & Sautiere, P. *Eur. J. Biochem.* **147**, 351–359 (1985).
2. Wouters-Tyrou, D., Sautiere, P. & Biserte, G. *FEBS Lett.* **65**, 225–228 (1976).
3. Sobel, R. E., Cook, R. G. & Allis, C. D. *J. Biol. Chem.* **269**, 18576–18582 (1994).
4. Folk, J. E. *Rev. Biochem.* **49**, 517–531 (1980).
5. Takao, T., Fukuda, H., Coull, J. & Shimonishi, Y. *Rapid Commun. Mass Spectrom.* **8**, 925–928 (1994).

Collisions and collusion

Jeremy Bernstein

Lise Meitner: A Life in Physics. By Ruth Lewin Sime. University of California Press: 1996. Pp. 512. \$30.

SOME half a century after the fact, the mosaic of the German wartime nuclear efforts — from the discovery of nuclear fission at the end of 1938 to the internment of ten leading German scientists at Farm Hall, near Cambridge, England, in the summer of 1945 — is beginning to fill itself in.

There are overviews of the programme (for example *German National Socialism and the Quest for Nuclear Power, 1939–1949* by Mark Walker; Cambridge University Press, 1989). There are biographies of some of the principals such as Werner Heisenberg (see, for example, *Uncertainty: The Life and Science of Werner Heisenberg* by David C. Cassidy; W. H. Freeman, 1992). There is now a completely annotated edition of the so-called Farm Hall transcripts, the recorded conversations of the interned physicists (see *Hitler's Uranium Club* by Jeremy Bernstein, with an introduction by David Cassidy; American Institute of Physics Press, 1996). There are still a few missing pieces, such as an objective biographical study of the murky von Weizsäcker family, which probably will not happen while Carl Friedrich von Weizsäcker is still alive.

But now we have another piece of the puzzle filled in with a detailed new biography of the only woman involved in the German uranium story: the Austrian-born physicist Lise Meitner. Written by the American chemist Ruth Lewin Sime, it is clearly a labour of dedication, if not of love; it seems to have taken her some 20 years to complete. It is also a partisan book. Sime clearly does not much like Meitner's German colleagues such as Heisenberg, von Weizsäcker and Meitner's life-long friend and co-worker Otto Hahn. As I don't much like them either, her point of view does not bother me at all, although it may trouble some readers.

Lise Meitner was born in Vienna in 1878 and lived there for 29 years. Her father was a lawyer and the family was of Jewish descent. But, in the decades after Meitner's birth, they all converted to Christianity — something that was very

common in middle- and upper-class Jewish Austro-Hungarian circles. Meitner found in the 1930s that the Nazis did recognize nuances such as baptisms when it came to the Jews. She was fortunate in the timing of her birth. By the turn of the century it was possible for a woman in Austria to receive a higher education at a university, even in science. Her teacher was the Austrian theoretical physicist Ludwig



Lise Meitner: friendship with Hahn was broken and betrayed.

Boltzmann. He must have been an astonishing lecturer. My own teacher, Philipp Frank, who was also a student of Boltzmann and who knew and had heard all the great physicists of the first half of this century, told me that Boltzmann was the best. But Boltzmann committed suicide in 1906. Meitner decided that the only future for her in physics was to move to Germany, which she did in 1907, and she remained in Berlin until she was forced to leave in 1938.

One of the first people she met there was the chemist Otto Hahn. They were

nearly the same age. (Actually Hahn was four months younger, important to remember when we turn to the aftermath of the discovery of fission.) They began a collegial relationship that lasted for more than 30 years. Sime does a competent job of describing Meitner's work. There is no space to go into the details here, suffice to say that what struck me was that Meitner was a very good physicist but not a great one. In a surprising number of instances she came in second. Having just missed a great discovery, she was then in an excellent position to confirm or expand on the new result. One might say that this was 'bad luck', but the really great experimental physicists such as Ernest Rutherford and Enrico Fermi make their own luck. Nonetheless, she was impressive enough that in 1917 she was charged with establishing a physics department at the Kaiser Wilhelm Institute in Berlin where Hahn had a chemistry department. The institute was not a university so for a while it could shelter 'non-Aryans' such as Meitner, who was, until the *Anschluss* in 1938, protected by her Austrian nationality. She waited too long to leave Germany and was finally forced to flee for her life, settling unhappily in Sweden.

Meanwhile, Hahn and his associate Fritz Strassmann — a courageous anti-Nazi — continued the work begun there by Meitner on the collisions of 'slow' neutrons with uranium. This programme had been invented by Fermi in Rome a few years earlier. (For a description and discussion of why the group failed to discover fission, see *Enrico Fermi, Physicist* by Emilio Segrè; University of Chicago Press, 1962 and 1995.) It could have led to the discovery of fission several years before 1938. The reason why it did not is that until the end of 1938, when Meitner and her nephew Otto Frisch realized what the physics was, the idea of a slow

neutron (the slower the better) breaking up a heavy nucleus such as uranium seemed absurd. What physicists, from Fermi onward, thought would happen is that the neutron would chip off a few nuclear particles from the uranium leaving behind another heavy element. Once the energetics were properly understood by Meitner and Frisch, the process was viewed as more like a gentle breeze blowing an unstable rock off a cliff. But it was only after the experiments of Hahn and Strassmann that all this changed. They found that the relatively light element

barium was left over after the collision — a finding they did not understand at all. Hahn appealed to Meitner in Sweden and she and Frisch created the theory.

There is no question that this discovery deserved a Nobel prize, but the actual award to Hahn alone, while he was in detention in England, is certainly one of the worst made by the Royal Swedish Academy of Sciences. There should have been a joint award to Hahn and Strassmann in chemistry and a joint award to Meitner and Frisch in physics. What the Swedish Academy thought it was doing I have no idea. Sime thinks the oversight may have had something to do with a personal animosity between Meitner and the Swedish physicist Manne Siegbahn. Knowing how these things work, this strikes me as entirely plausible. Hahn's reaction to all of this may be humanly understandable, but it is despicable nonetheless. He gave Strassmann a sort of 'tip' — 10 per cent of his prize money. And then, like the other German nuclear physicists, he began to rewrite history. This began at Farm Hall. On 8 August 1945, the ten Germans signed a document giving their version of their wartime activities. There is a 'remark', one of several attached to the body of the document, that reads: "The Hahn [*sic*] discovery was checked in many laboratories, particularly in the United States, shortly after publication. Various research workers — Meitner and Frisch were probably [*sic*] the first — pointed out the enormous energies which were released by the fission of uranium. On the other hand, Meitner had left Berlin six months before the discovery and was not concerned herself in the discovery." There is no charitable way to read this abominable statement. The Germans might at least have had the decency to explain why Meitner had left Berlin.

Although I am generally admiring of Sime's book, I do have a few criticisms. Here are three. First, she has fallen into what I call the Kramish trap. In 1986 Arnold Kramish wrote a book about the successful espionage efforts of Paul Rosbaud, an Austrian chemist who remained in Germany (*The Griffin: The Greatest Untold Espionage Story of World War II*; Houghton Mifflin, 1986). Kramish quotes an extraordinary letter from Max von Laue (an anti-Nazi German physicist who was interned at Farm Hall) to Rosbaud. In it von Laue describes how the internees created the '*lesart*' (or version) of their activities, which they have been propagating ever since. Laue writes: "The leader in these discussions was Weizsäcker. I did not hear the mention of any ethical point of view. Heisenberg was mostly silent." In Kramish's book, and in Sime's, the last sentence is underlined and the underlining is attributed to Laue. But copies of the original letter show no underlining. Second, contrary to what Sime says, I do

not believe that Hahn was awarded the Nobel prize in 1944 and so informed. The Farm Hall transcripts, and other contemporary documents make it clear that Hahn was flabbergasted when his prize was announced in 1945. Finally, and most importantly, I take issue with Sime's observation about the Farm Hall transcripts themselves. She writes: "The Farm Hall transcripts... do not shed much new light on the tangled issues of competence and intent". Not to put too fine a point on it, I think this statement is nonsense. My quotations from the transcripts in this review may make that clear. But above all, the transcripts settle, once and for all, the question of what the Germans knew about nuclear bomb physics. They knew next to nothing. That being so, the idea that for moral reasons they did not 'give' the bomb to Hitler falls of its own weight. □

Jeremy Bernstein is at the Aspen Center for Physics, 600 West Gillespie Street, Aspen, Colorado 81612, USA.

Chemical contexts

Philip Ball

The Same and Not the Same. By Roald Hoffmann. Columbia University Press: 1995. Pp. 294. \$29.95.

HERE is a story Roald Hoffman might appreciate. A friend was dining recently in a London restaurant, and asked for a bottle of non-sparkling mineral water. A bottle arrived at the table — unlabelled, covered in condensation and effervescing gaily (it had been opened already). My friend pointed out that it was not what he had ordered, and the bottle was removed and replaced by still water in another, apparently identical but opened bottle. Suspicious, he asked for an unopened bottle. "We can't do that", said the waiter. "We make our own mineral water here" — by chilling and carbonating water from the tap. (Presumably the still version was produced by allowing the carbonated water to degas?)

Hoffmann's book dwells on the difference between the 'real' and the 'synthetic' — a difference we see increasingly advertised in foods, health-care products and domestic reagents. It is this duality that underlies so much of the public perception of chemistry. One might expect, from a chemist, a scornful tirade about these apparently meaningless distinctions — does not 'natural' vitamin C extracted from rosehips have the same molecular structure as 'synthetic', which started life as crude oil? But instead, Hoffmann's response is sympathetic: "That a reasonable human being can be ambivalent about chemicals... is not a sign of irra-

tionality but of humanity." To arrive at this conciliatory position, he does not have to relinquish the chemist's precision: he points out that the differences between 'rosehip' and 'synthetic' vitamin C, for example, are not some vitalistic figment of a health freak's imagination but the inevitable consequence of product impurities at the parts-per-thousand level. Whether those differences matter or not is another thing entirely.

But primarily, Hoffmann wants to tell readers not why chemistry is important but why he loves it. And he understands that this requires a thumbnail sketch not of what chemistry is but of what chemists really do. The two are (the same and?) not the same. One might consider the job well done if a book about chemistry talked about its triumphs, its ingenuity, its relevance to our lives. But to paint a picture of the craft (and not just list its achievements), the book should also talk about chemistry's failures and disasters, its practitioners, and its modes of representation and communication.

That all of this applies to any scientific or intellectual pursuit goes without saying. But it is not often that writers of popular science pay it more than lip service. And so lay readers are left with the conundrum (if they were to delve this deep) that the fantastic new discoveries breathlessly recounted between glossy covers are the very same as those reported in the scientific literature in a stilted, stylized and recondite manner of Victorian and almost wilful dourness. Hoffmann delves into this literature to reveal the many-layered, barely conscious meanings struggling within the straitjacket of conventional scientific prose. I am, however, a little sad (and a little surprised) that so eloquent and thoughtful a commentator does not rail more against the austere absurdities that the scientific literature approves. Probably he can ignore them in his own papers, and assumes that others will do likewise.

Then there is the ritual of publication itself — surely a bizarre process to the uninitiated. An editor at *Nature* cannot read Hoffmann's extracts from referees' reports without a wry smile (nor help but wonder how selective these extracts have been). "Hoffmann is very intelligent — but not intelligent enough to do anything positive", fumes a reviewer for the *Journal of the American Chemical Society*. "You chemists should raise your standards", declares a physicist. Hoffmann feigns baffled surprise at these outbursts, but I suspect he knows that they represent nothing more than people letting off steam in print, and that surprise or expectations of rationality have no part of it.

But not everything Hoffmann says about the pursuit of chemistry can be generalized. In his description of the controversial, often unhappy life of Fritz Haber,

or his account of the thalidomide disaster, there is ample indication that chemistry impinges on life in ways that, say, particle physics does not. He doesn't actually say this, and to suggest it here will probably provoke outrage; but what is the point of disciplinary relativism? When you are dealing with tangible stuff rather than ideas or theories, and particularly when you are selling it to the public, errors and oversights (and, in the case of thalidomide, deliberate obfuscation) can be serious in a way that a faulty integral is not. That's another reason why I like this book — it weaves chemistry into the evolution of a technological society, making this in part a book about the sensible world rather than about the abstract hypotheses that sell so much popular science. It dips tantalizingly into the untold histories of twentieth-century science, the ones not told because they do not fit the heroic ideal of science's striving for pure knowledge: of I. G. Farben and the German dye industry, of pharmaceuticals and agrochemicals, of the art of making things.

All this is a way of saying that I should be surprised if any scientist did not enjoy this book. If some find Hoffmann too liberal in his tolerance of chemistry's 'green' critics, perhaps they might ask themselves to whom those critics would give the more sympathetic hearing. Hoffmann's remarks on the limitations of reductionism and the scientist's susceptibility to the myth of rationality in dealing with the world of human affairs will win him critics. But who can argue, when a supremely rational demolition of astrology by the zoologist Richard Dawkins in a 'quality' British newspaper engenders not converts to reason but one of their biggest and most outraged mail-bags? Winning arguments in public and in politics is not about the primacy of fact, suggests Hoffmann, but about the recognition of (and concession to) irrational prejudice and belief. Not so different from the peer-review process, in fact. □

Philip Ball is an associate editor of *Nature*.

Natural products

Wiley has just published the second edition of *Encyclopedia of Common Natural Ingredients Used in Food, Drugs, and Cosmetics* by Albert Y. Leung and Steven Foster. Encompassing some 500 naturally derived ingredients, this substantially revised and updated edition addresses the increased interest in natural ingredients over the past decade, especially in health foods — a category that includes food products, over-the-counter drugs and herbal teas. Also comprehensively covered are Chinese medicinal herbs. Price is \$175, £115.

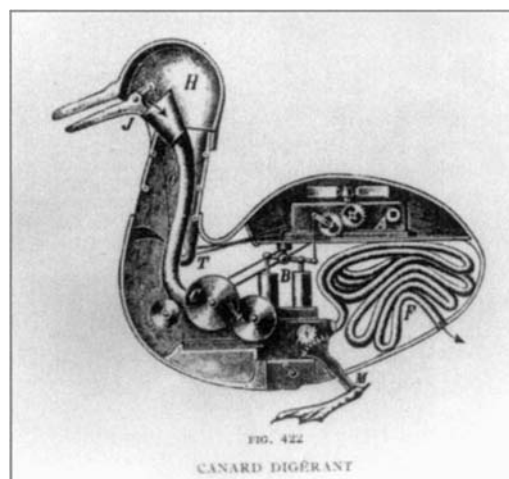
Truth in the balance

Willem Hackmann

The Invisible World: Early Modern Philosophy and the Invention of the Microscope. By Catherine Wilson. Princeton University Press: 1995. Pp. 280. \$39.50, £32.

Instruments and the Imagination. By Thomas L. Hankins and Robert J. Silverman. Princeton University Press: 1995. Pp. 337. \$39.50, £33.50.

WHAT do the sunflower clock, microscope, air pump, tuning-fork synthesizer and mechanical duck have in common? They were all considered in their day to be



Clockwork quackery: Vaucanson's mechanical duck.

scientific instruments. In perusing such a diverse list, one thing becomes immediately apparent, namely that the concept of 'scientific instrument' has covered a very broad church indeed since the start of modern science in the sixteenth century. Perhaps it is no accident that the authors of these two books are respectively a philosopher, historian and lawyer. Despite their different titles, both books deal with the power of instruments in stimulating the intellect and the imagination. *The Invisible World* focuses on the influence of the microscope which, in the seventeenth century, opened up a new world of observation, whereas *Instruments and the Imagination* demonstrates that instruments not only could be at the cutting edge of research, but were also often crucial in defining the boundaries of acceptable science. The second book is full of interesting illustrations, whereas the first — conforming with its title — has none. It is somewhat hard to imagine a book on the influence of the microscope with no pictures, but then this is very much a philosophical treatise (perhaps philosophers think but do not see).

Recently, there has been a growing

interest in the role of instruments in the study of nature, both by historians and philosophers of science, and even by a sprinkling of art historians fascinated by the images produced by such devices. These two books demonstrate that we have moved a long way from an earlier generation of historians of science (chiefly composed of retired or semi-retired scientists) who regarded instruments as mere gadgets and voiced a commonly held prejudice that "The love of gadgets is only the beginning of wisdom" (H. T. Pledge in *Science since 1500*; HMSO, 1966).

Mechanical means of recreating nature, human bodily functions included, have intrigued people since antiquity. What could be more extraordinary than replicating the human voice? Stories abound of ancient speaking statues, such as the Greek head of Orpheus at Lesbos, but the

effects were produced by concealed priests speaking through tubes or by ventriloquism. Fascination with automata, including imitation of the voice, was an important aspect of the circle of natural philosophers who eventually founded the Royal Society in 1660. Particularly notorious was John Wilkins's speaking statue, which Wilkins kept in his garden for playing tricks on his friends. But this, too, was operated by way of a long concealed tube. Jacques Vaucanson's mechanical duck of 1731, however, was far more sophisticated, for by then advances in clockwork technology, in particular in miniaturization, had made complicated automata possible.

The duck could eat, drink and (lo and behold) excrete, thereby siding with those who argued that such bodily functions were purely mechanical and required no divine intervention. But Vaucanson also appreciated that money could be made from the entertainment value of his mechanical production. The duck thus neatly showed two important characteristics of science, serving as both a public spectacle and a new way of describing the world. As Thomas Hankins and Robert Silverman point out, such attempts to create an artificial world in what we now call early modern science may well seem pointless to us, but imitating nature in fact offered the possibility of acquiring as well as demonstrating an understanding of phenomena. Indeed, is not much of our understanding of the world today based on model experiments and analogous argument? Perhaps science is no more than a sophisticated metaphor for the world we inhabit.

That German polymath Hermann von Helmholtz in a lecture in 1854 understood the significance of Vaucanson's duck precisely because to his generation of scientists it was regarded as a misdirected



Blooming marvellous: Kircher's sunflower clock.

effort towards a mechanical description of nature. According to Hankins and Silverman, even if one assumes that Helmholtz shared Vaucanson's goal, there remains an obvious qualitative difference between the mechanical duck and Helmholtz's tuning-fork synthesizer. Vaucanson's duck copied animal functions as they were observed (from the outside); Helmholtz's

apparatus copied what he thought to be the mechanism of sound. From our perspective, Helmholtz's approach is the more scientific. A modern comparison perhaps is between the old way of creating a flavour by mixing chemicals and the modern way by copying the molecule responsible for that flavour.

Athanasius Kircher's sunflower clock which purported to tell the time through a cosmic magnetic force teaches us the same lesson. By our standards the clock was a fraud but definitely not by his. No-one can deny that there is an ocean of difference between the sunflower clock and the microscope or the air pump, but in a real sense all three were emblems of their time: the sunflower clock of early-seventeenth-century natural magic, the microscope of the new generation of sense-expanding instruments of mid-seventeenth-century natural philosophy and the air pump of the new generation of laboratory devices of experimental philosophy that allowed active exploration of natural phenomena in the laboratory.

Hankins and Silverman ask the important question whether the inability of Helmholtz's generation to understand Vaucanson's duck (and by the same token Kircher's sunflower clock) was caused by a shift in the boundaries of acceptable sci-

ence, or by a completely different concept of what science should be. The fact that the question can be asked suggests that the criteria for what counts as 'scientific' have never been fixed. That is the core message of their book.

Catherine Wilson's treatise on the impact of the microscope on seventeenth-century thought covers very much the same ground. Her key years are the period from 1620 to about 1720, by which time the main arguments about the validity of observations made with this device had been resolved.

Microscopists had to deal with such knotty issues as the relative status of fact and observation; and the historian evaluating their work must decide whether seventeenth-century 'proto-science' was an immature form of modern science or something else altogether — a similar problem to that faced by Hankins and Silverman over Helmholtz. Wilson shows that microscopic observations reinforced the contemporary idea of the 'living machine' — that is, a reductionist view of nature. And therein lies the ultimate paradox of our machine-driven science, which is highlighted by both books: the essence of our natural world remains hidden despite our increasingly sophisticated scientific technology. Perhaps *The Invisible World* is better off with no pictures after all. □

Willem Hackmann is at the Museum of the History of Science, University of Oxford, Broad Street, Oxford OX1 3AZ, UK.

Underground art

ON 18 December 1994, three cavers in the Ardèche, southeast France, stumbled across more than 300 wall paintings dating back 30,000 years in a hidden underground cavern — the world's oldest examples of prehistoric art and some 15,000 years older than the renowned paintings at Lascaux. Chauvet cave is startling in many respects: it is completely intact, showing varied traces of prehistoric human and animal visitors, including the already notorious bear skull placed on a rock, which has led to much untrammelled speculation about bear cults; and it is decorated with images of many different species — rhinoceroses, big cats and bears — previously unknown in the region's Ice Age art, but also rare elsewhere, and certainly not usually depicted with such prominence. For example, fewer than 20 rhinos were previously known in European cave art, whereas Chauvet contains between two to three times that many alone. The images are particularly impressive for the techniques used to present perspective and motion: many figures interact, some are staggered and others are drawn on bulges in the cave walls to suggest depth. Unlike Lascaux, this cave is now closed to the public and its priceless heritage has been seen by only a privileged few. In *Chauvet Cave*, Jean-Marie Chauvet, Eliette Brunel Deschamps and Christian Hillaire now recount the tale of their discovery and present colour photographs of the remarkably sophisticated paintings, which are analysed and set in context by the prehistoric art specialists Paul Bahn and Jean Clottes. Thames and Hudson, £28.

P. T.



A common plan for dorsoventral patterning in Bilateria

E. M. De Robertis & Yoshiki Sasai

Functional studies seem now to confirm, as first suggested by E. Geoffroy Saint-Hilaire in 1822, that there was an inversion of the dorsoventral axis during animal evolution. A conserved system of extracellular signals provides positional information for the allocation of embryonic cells to specific tissue types both in *Drosophila* and vertebrates; the ventral region of *Drosophila* is homologous to the dorsal side of the vertebrate. Developmental studies are now revealing some of the characteristics of the ancestral animal that gave rise to the arthropod and mammalian lineages, for which we propose the name *Urbilateria*.

One of the challenges in evolutionary biology is to ascertain to what extent animals have homologous structures that have been derived by descent from a common ancestor (as in the case of the pectoral fin of fishes, forelimbs of tetrapods, and wings of birds and bats) or whether any similarities are due to convergent evolution resulting from the need to perform similar functions (as in the case of wings of flies and birds). Fifty years before Darwin, French naturalist E. Geoffroy Saint-Hilaire proposed that the ventral side of the arthropods was homologous to the dorsal side of the vertebrates¹. He dissected a lobster (Fig. 1), but instead of placing it in its usual orientation with respect to the ground, he placed it upside down. In this orientation the lobster's central nervous system (CNS) was located above the digestive tract, which in turn was located above the heart. In his own words (our translation): "What was my surprise, and I add, my admiration, in perceiving an ordering that placed under my eyes all the organic systems of this lobster in the order in which they are arranged in mammals?" This idea of a dorsoventral inversion between arthropods and mammals led to a dispute with Georges Cuvier² in the French Academy, and has been revived³⁻⁶ and contested^{3,7-9} several times. Here we re-examine the issue of homology versus convergence in the development of animal species from the perspective of recent results in molecular embryology.

Dorsoventral patterning by *chd* and *sog*

In vertebrates the formation of the dorsal mesoderm and of the CNS are induced by a region of the embryo called the organizer. *Xenopus chordin* (*chd*) is an organizer-specific secreted protein that can mimic the activity of the organizer, resulting in a twinned body axis¹⁰. A *Drosophila* complementary DNA of related structure (Fig. 2), was reported^{11,12} and, interestingly, it encoded *short gastrulation* (*sog*), a gene long known to play an important role in *Drosophila* dorsoventral patterning^{13,14}. As shown in Fig. 3, a number of zygotic genes are required for dorsoventral patterning of the *Drosophila* embryo. The regions in which these genes are expressed in the embryo are controlled by the activity of the maternal gene *dorsal*, which can repress or activate their transcription. Most are expressed in the dorsal 40% of the embryo, and are required to produce active *decapentaplegic* (*dpp*) gene product, which is a growth factor of the transforming growth factor- β (TGF- β) family. *Dpp* diffuses, generating a morphogen gradient that patterns the ectoderm^{15,16}. Conversely, *sog* is expressed ventrally and its protein diffuses dorsally, where it antagonizes the activity of *dpp* (Fig. 3).

When the embryonic expression patterns of *chd* in the vertebrate¹⁰ and *sog* in the fly¹¹ are compared, it can be observed that they are inverted with respect to each other (Fig. 4). *sog* is initially expressed in the ventral 60% of the blastoderm, and by the end of

gastrulation it becomes localized to two rows of cells, called the mesectodermal cells, along the ventral midline. *chd* is expressed initially in the dorsal side of the blastopore (that is, in the organizer region), and by the end of gastrulation it is restricted to the dorsal midline (Fig. 4).

The case of *sog* and *chd* allows us to test whether ventral in the fly corresponds to dorsal in the vertebrate, because the gene products (in the form of messenger RNAs) have biological activity when injected into embryos. Injected *sog* mRNA ventralizes fly embryos, leading to formation of ventral denticle belts and ectopic patches of CNS¹⁷. In *Xenopus*, *sog* mRNA causes dorsal development (for example, formation of dorsal tissues such as notochord and CNS)¹⁷. The vertebrate homologue, *chd*, behaves functionally much in the same way as does *sog*: injection of *chd* mRNA, a dorsalizing factor in *Xenopus*¹⁰, promotes ventralization of cell fates in *Drosophila*¹⁷. Thus, the function of *sog* and *chd* is reversed in insects and vertebrates; in both cases injection of the gene product promotes the development of the side of the embryo that contains the CNS.

dpp and *Bmp-4* antagonize *sog/chd*

In *Drosophila*, *dpp* is expressed dorsally and promotes dorsal development^{15,16}, that is, it is the opposite of *sog*. Microinjection of increasing amounts of *dpp* mRNA leads to threshold responses in the pattern of tissue differentiation in *Drosophila* ectoderm, shifting the balance in the dorsal direction until at high doses ventral development is prevented^{15,16}. In the vertebrate, *Bmp-4* (bone morphogenetic protein 4) and *Bmp-2* share extensive sequence similarity with *Dpp*. In the *Xenopus* gastrula *Bmp-4* is expressed in the ventral and lateral marginal zone (which gives rise to ventral mesoderm) as well as in the animal cap (which gives rise to skin ectoderm when cultured in explants)¹⁸. In *Xenopus* embryos *Bmp-4* has potent activity, changing the fate of dorsal mesoderm into ventral cell fates (blood and mesenchyme)^{19,20}. *Drosophila dpp* can mimic the ventralizing activity of *Bmp-4* in *Xenopus* and this activity can be antagonized by injected *sog* mRNA¹⁷. The gradient of *dpp* activity observed in *Drosophila* may result, at least in part, from the antagonism between the *Sog* and *Dpp* diffusible factors²¹. It is not yet known whether *Sog* antagonizes by binding directly to *Dpp* or its receptor, or whether it acts through its own receptor. A similar antagonism has been reported for *chd* and *Bmp-4* (ref. 22).

The suggestion that *Bmp-4* is the vertebrate homologue of *dpp* sparked the recent interest in the inversion of the dorsoventral axis⁴. The functional experiments discussed above indicate that dorsoventral patterning in both *Drosophila* and *Xenopus* is dependent upon a system of antagonistic extracellular signals provided by *dpp/Bmp-4* and *sog/chd*. Despite the morphological differences

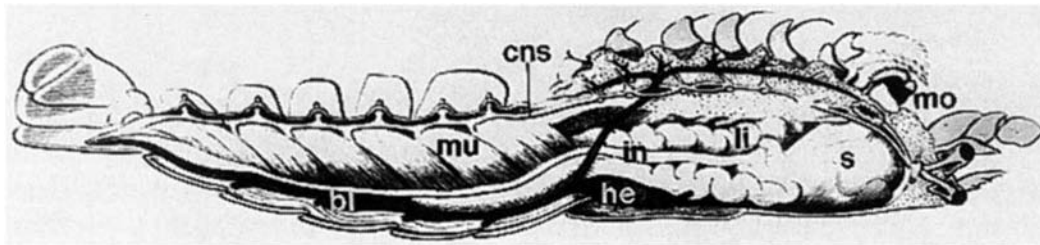


FIG. 1 Geoffroy Saint-Hilaire's famous lobster. In this dissection the animal is presented in the orientation opposite to that it would normally have with respect to the ground. The central nervous system (cns or nerve cord) is above, and is traversed by the mouth (mo). Underneath is the digestive tract, with the stomach (s), liver (li) and intestine (in). Below the gut are the heart (he) and main blood vessels (bl). Muscles (mu) flank the CNS. In this orientation the body plan of the arthropods resembles that of the vertebrate. From ref. 1, by courtesy of the History and Special Collection Division, Louise M. Darling Biomedical Library, UCLA.

between embryos of the two species, the *sog/chd* gene is expressed on the side from which the CNS arises while the *dpp/Bmp-4* gene is expressed on the opposite side of the embryo. The functional conservation of the *sog/chd* and the *dpp/Bmp-4* secreted proteins suggests a homologous mechanism of dorsoventral patterning that must have existed in the common ancestor from which insects and vertebrates diverged. The results support the view that a reversal of the dorsoventral axis occurred during the course of evolution.

The inversion challenged

The recent revival of Geoffroy Saint-Hilaire's inversion hypothesis⁴ was opposed⁷⁻⁹ and this merits some discussion in light of recent findings. Three main objections have been raised.

First, it has been argued that the use of marker genes, such as those of the *achaete-scute* complex, that are conserved in vertebrates and insects⁴ is inappropriate because they mark the presence of similar tissues, in this case nerve cells, and not homologous topology^{7,8}. This objection does not apply to the case of *sog/chd* and *dpp/Bmp-4* because these molecules are

upstream regulators providing patterning information that determines tissue differentiation according to specific dorsoventral fates. The importance of identifying the upstream regulatory switches is highlighted by the recent discovery that eye development is controlled by a conserved gene, *Pax-6/eyeless*, in both vertebrates and *Drosophila*^{23,24}. This is a good example because previously the formation of eyes in insects and mammals was one of the classic cases of convergent evolution, despite the presence of common molecular elements in the light-detection system, such as conserved opsin proteins²⁵. Because eye development is controlled by the same genetic switch in flies and mice it is now considered that eyes are homologous structures related by descent from an ancestral photoreceptor²⁶.

Second, the homology between *dpp* and *Bmp-4* has been challenged on the basis of the similarities between *Bmp-4* and its close relative *Bmp-2* (refs 7, 8). Indeed, *Bmp-2* and *Bmp-4* are very similar in sequence, and both can functionally substitute for *dpp* in *Drosophila*²⁷. However, recent loss-of-function studies

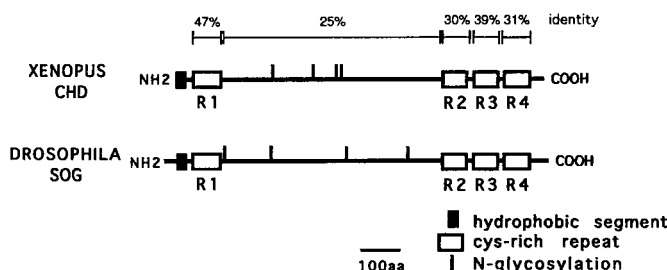


FIG. 2 The amino-acid sequences of *Xenopus chd* and *Drosophila sog* share similarities. Both proteins have a secretory signal sequence, or hydrophobic segment, at the amino (NH2) end (dark box), several putative N-glycosylation sites (vertical lines) and four cysteine-rich repeats. The spacing of the nine cysteines contained in the cys-rich repeats is conserved in other secreted proteins such as thrombospondin, von Willebrand factor and procollagen. The first repeat (R1) of *chd* is more similar to R1 of *sog* than to any of the other repeats in *chd*. The same is true for R4, indicating that both genes are derived from a common ancestor. The percentage of amino-acid identity between the two proteins is indicated. The amino-acid identity is only 28% along the entire protein, so it was especially important to show that *sog* and *chd* were functionally homologous. Because of its homology of *sog*, and for other reasons⁴⁹, we propose the designation *schordin* to refer to the vertebrate homologue (*chd*) in future.

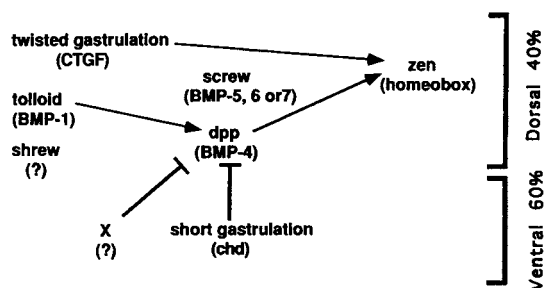


FIG. 3 Zygotic genes involved in dorsoventral patterning of the ectoderm in *Drosophila* embryos, and their vertebrate relatives. Several genes are expressed in the dorsal 40% of the embryo; *tolloid* encodes a metalloprotease required for *dpp* activity, and is similar in sequence to vertebrate *Bmp-1*; *twisted-gastrulation* encodes a protein related to vertebrate connective tissue growth factor; *screw* is a secreted growth factor related to either *Bmp-5*, 6 or 7 that might form heterodimers with *dpp*; *zen* encodes a homeobox gene for which a vertebrate counterpart has not been isolated. Injection of *short-gastrulation* (*sog*) product in wild-type *Drosophila* embryos leads to formation of ectopic ventral tissues and CNS, but in embryos mutant for *dorsal* it can only produce a partial ventralization¹⁷; this led to the proposal that a second ventralizing gene (indicated as 'X'), that cooperates with *sog* and is activated by *dorsal*, should also exist¹⁷.

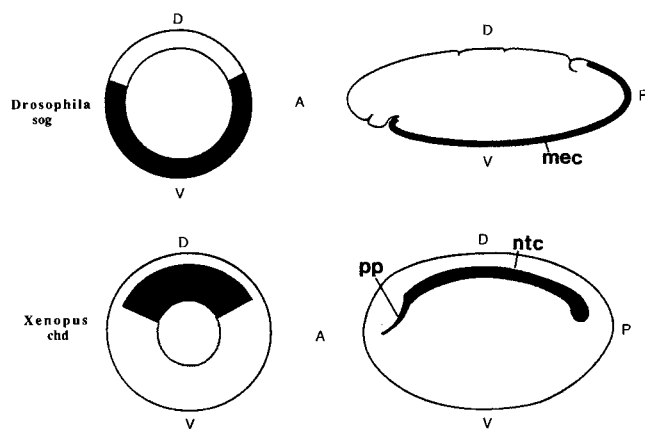


FIG. 4 The expression of *Drosophila sog* (top) and *Xenopus chd* (bottom) is reversed with respect to the dorsoventral axis. In *Drosophila sog* is first activated in the ventral 60% of the blastoderm, is then switched off in the mesoderm and persists in the neurogenic ectoderm, and by late gastrulation the expression resolves into two rows of cells located in the ventral midline, the mesectodermal cells (*mec*)¹¹. In *Xenopus*, *chd* is expressed initially in the dorsal side of the blastopore, whereas by the end of gastrulation it is resolved to the mesoderm of the dorsal midline, occupying the head mesoderm or prechordal plate (*pp*) and the notochord (*ntc*)¹⁰. *sog* is ventral in *Drosophila* whereas *chd* is dorsal in *Xenopus*.

using antisense RNA in *Xenopus*²⁸ and gene disruption in mice^{5,29} indicate that Bmp-4, but not Bmp-2, plays an essential role in early dorsoventral patterning of the vertebrates, supporting the view⁴ that Bmp-4, rather than Bmp-2, is the homologue of *dpp* that functions in early embryos.

Third, it has been argued that the displacement of the CNS from ventral to dorsal could have resulted from the appearance of a novel region of ectoderm around the mouth region⁷. In this view, first proposed by Garstang, the region corresponding to the ventral midline of the ventral CNS would become displaced, forming the lateral edges of the dorsal neural plate in vertebrates⁷. Recent molecular findings strongly argue that this is not the case, because the midline of the CNS in *Drosophila* and in the vertebrate share a conserved signalling molecule called *netrin*. The vertebrate *netrin-1* gene encodes a secreted protein expressed in the midline of the CNS (the floor plate) that guides commissural axons to the opposite side of the developing spinal cord³⁰. A *netrin* homologue is also expressed in the midline of the *Drosophila* CNS (C. Goodman, personal communication) and in the *Caenorhabditis elegans* ventral nerve cord³¹. Similarly, in the case of enteropneusts^{8,9} and other vermiform animals with multiple nerve cords the region homologous to the CNS will ultimately have to be defined molecularly by the expression of *chd* and *netrin* homologues.

Thus, although arguments opposing the dorso-ventral inversion have been raised, the recent results from molecular biology tend to vindicate the hypothesis of Geoffroy Saint-Hilaire.

Common signals

An unexpected conclusion from recent studies on dorsoventral patterning is that the ectoderm and the mesoderm appear to be patterned by a common set of signalling molecules involving *dpp*/*Bmp-4* and *sog/chd*^{22,32}. In *Drosophila*, there is strong genetic evidence that the amount of ectoderm allocated to the neurogenic region is controlled by positional information provided by this system. Both loss-of-function and gain-of-function studies show that *dpp* has an antineurogenic activity^{15,16,21}, promoting the formation of dorsal epidermis, whereas *sog* promotes neurogenesis in the ectoderm^{11,17,21}. *Dpp* is widely recognized as the main morphogen patterning the ectoderm, but recent evidence indicates that *dpp* patterns the *Drosophila* mesoderm as well. The mesoderm expresses a Dpp receptor, *thickveins*³³, and Dpp produced by the

ectoderm activates the expression of dorsal mesodermal genes (*bagpipe* and *tinman*) and represses the expression of a ventral mesodermal marker (*pox meso*)^{34,35}.

In *Xenopus*, CNS formation is thought to be different to *Drosophila*. The vertebrate neural plate is considered to be induced by signals secreted by the organizer. The organizer is also considered to be responsible for the release of signals that promote dorsal differentiation of mesoderm³⁶. It was generally assumed that these signals would be different, but it now appears that the same molecules that pattern mesoderm also regulate CNS formation, as depicted in Fig. 5. Microinjection of *chd* (or *sog*) mRNA changes the fate of gastrula animal cap explants from ventral ectoderm (epidermis) to dorsal ectoderm (neural tissue)²². In ventral marginal zone explants (which form ventral mesoderm such as blood and mesenchyme) *chd* induces dorsal mesoderm (notochord and muscle)¹⁰. A similar situation has been observed for two other organizer-specific secreted factors, noggin and follistatin, which are neural inducers as well as dorsalizing agents^{22,37-39}. This indicates that the signalling molecules used for dorsal differentiation of both ectoderm and mesoderm are the same, and that the differences must reside in the responding tissues (Fig. 5).

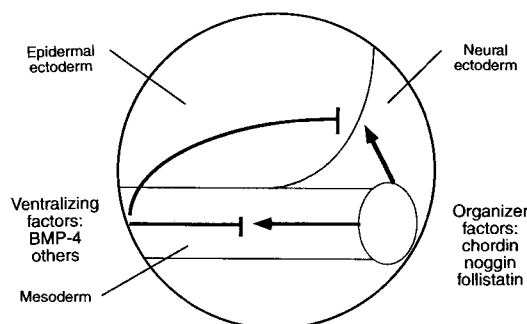


FIG. 5 Model indicating that the same set of regulatory signals may provide the positional information that patterns both ectoderm and mesoderm in *Xenopus*. On the dorsal side (right), the organizer (oval) provides dorsal positional values to ectodermal (animal cap, top) and mesodermal (marginal zone, middle) tissues by secreting organizer factors such as *chd*, *noggin* and *follistatin*. On the opposite side (left), ventralizing factors such as Bmp-4 and presumably other signals give ventral positional values to the tissues, antagonizing organizer signals. High dorsal values promote neural differentiation in the ectoderm and formation of notochord and muscle in mesoderm, whereas high ventral values lead to epidermogenesis in ectoderm and the differentiation of blood island and mesenchyme in the mesoderm. In this model, a default status of tissues does not exist and the balance between the dorsal and ventral signals decides dorsoventral fates.

Studies on animal cap explants suggest that Bmp-4 has potent antineurogenic effects in *Xenopus*. Animal cap cells contain endogenous Bmp-4 mRNA¹⁸ which is required to prevent neural differentiation²². In addition, a recent experiment suggests that the neuralizing effect of cell dissociation on animal caps^{40,41} may be due to the dilution of the Bmp-4 signal by diffusion into the culture medium. Indeed, when Bmp-4 was added to dissociated cells, they differentiated into epidermis (ventral ectoderm) and neural differentiation was prevented³². Overexpression of Bmp-4 can suppress neural induction caused by *chd* and, intriguingly, also neuralization by *noggin* and *follistatin*²². The antagonism between Bmp-4 and *chd* in neural induction is consistent with the roles of their homologues, *dpp* and *sog*, in *Drosophila* ectodermal patterning. It is not known whether *Drosophila* has homologues for *noggin* and *follistatin*, but this is an attractive area for future research because it has been proposed that additional factors should

cooperate with *sog* in *Drosophila* neurogenesis (see 'X' in Fig. 3)¹⁷.

Although *Drosophila* neurogenesis and *Xenopus* neural induction have traditionally been considered to be different, the molecular mechanisms underlying the patterning of two distinct germ layers, ectoderm and mesoderm, seem to be regulated by the conserved *sog/chd* and *dpp/Bmp-4* system. The concept of dorso-ventral patterning by positional information, which has been so useful in *Drosophila*, might also be applicable to vertebrates. In this view, the organizer would be the source of dorsal positional values which are counteracted by ventral values provided by *Bmp-4* and presumably other ventralizing factors (Fig. 5). Thus, the unity of plan revealed by the dorsoventral axial reversal issue can be extended to detailed tissue patterns within individual germ layers.

The Urbilateria or common ancestor

In 1874 Ernst Haeckel proposed a sweeping homology: that the ectoderm and endoderm in all metazoans was related by descent from a hypothetical animal called the *Gastrea*⁴². The *Gastrea* consisted of two cell layers, the ectoderm and the endoderm, forming a primitive gut cavity opening to the outside. Although simplistic, the *Gastrea* theory historically was very useful because it proposed that all multicellular animals were monophyletic.

A large amount of data have now become available on genes controlling development from *Drosophila* and the vertebrates. The hypothetical ancestral animal, for which we propose the name *Urbilateria* (primitive bilateral animal), from which the arthropod and the chordate lineages diverged 600 million years ago may have presented the following characteristics. (1) It had anteroposterior polarity determined by the Hox gene complexes^{43,44}. Additional genes present in *Drosophila* and vertebrates, for example homologues of *orthodenticle*, *empty spiracles* and *caudal* formed a network that cooperated with Hox genes in the generation of anteroposterior pattern⁴⁵. (2) It had a conserved system of dorso-ventral patterning provided by the antagonistic *sog/chd* and *dpp/Bmp-4* extracellular signals, as described above. (3) It presumably had a subepidermal longitudinal CNS that had well defined mid-line structures secreting axon guidance molecules such as netrin-1. (4) The function of *Pax-6/eyeless*^{23,24}, as well as the conservation of opsins²⁵, suggests that the common ancestor had primitive photo-

receptor cells from which the eyes of *Drosophila* and vertebrates evolved. (5) It presumably also had a circulatory system with a contractile blood vessel. *Drosophila* has a contractile dorsal vessel whose differentiation is controlled the homeobox gene *tinman* as well as by the transcription factor DMEF2, both of which have vertebrate homologues expressed in the developing heart tissue⁴⁶. In addition, other characteristics for which present evidence suggests independent evolutionary origins and convergent evolutionary solutions in insects and vertebrates, such as segmentation (metamerism) and the formation of appendages, may require revisions if new upstream regulatory systems are identified in future.

An important question is whether all deuterostomes evolved from the same ancestral animal. Evidence from 18S ribosomal RNA sequence comparisons is consistent with a common grouping for all deuterostomes⁴⁷. Thus it is possible that the inversion event may have occurred concomitantly with a great innovation in gastrulation mechanisms that led to the two main groups of metazoans: the protostomes and the deuterostomes^{4,48}. Most of the evidence on dorsoventral inversion of the axis reviewed here comes from comparisons of developmental genes in two very distantly related groups of organisms, the arthropods and the vertebrates. In future it will be useful to extend the comparisons of the *sog/chd* and *dpp/Bmp-4* signalling systems to other phyla in order to fill in the gaps, as has been so fruitful in the case of the Hox genes⁴⁴.

Note added in proof: Since this manuscript was first submitted, several papers have appeared that further substantiate the idea of a conserved ventral specification pathway in vertebrates. These include: (1) the expression patterns of *Bmp-4* and *Bmp-2* in *Xenopus*⁵⁰⁻⁵²; (2) additional support for the requirement of *Bmp* signalling to repress neutralization of animal cap explants⁵³⁻⁵⁵; (3) a role for *Bmp-7*, in collaboration with *Bmp-4*, as a ventralizing factor⁵⁵; (4) evidence for an inhibitory interaction between the ventralizing factor *Bmp-7* and follistatin; and (5) further evidence for the dorsalizing activity of *sog* in *Xenopus* embryos⁵⁷. □

E. M. De Robertis and Yoshiki Sasai are at the Howard Hughes Medical Institute, Department of Biological Chemistry, University of California, Los Angeles, California 90095-1737, USA.

- Geoffroy Saint-Hilaire, E. *Mém. du Mus. Hist. Nat.* **9**, 89–119 (1822).
- Appel, T. A. *The Cuvier-Geoffroy Debate* (Oxford University Press, Oxford, 1987).
- Nübler-Jung, K. & Arendt, D. *Wilhelm Roux Arch. dev. Biol.* **203**, 357–366 (1994).
- Arendt, D. & Nübler-Jung, K. *Nature* **371**, 26 (1994).
- Hogen, B. L. M. *Nature* **376**, 210–211 (1995).
- Jones, C. M. & Smith, J. C. *Current Biol.* **5**, 574–576 (1995).
- Lacalli, T. C. *Nature* **373**, 110–111 (1995).
- Peterson, K. J. *Nature* **373**, 111–112 (1995).
- Jefferies, R. P. S. & Brown, N. A. *Nature* **374**, 22 (1995).
- Sasai, Y. et al. *Cell* **79**, 779–790 (1994).
- François V., Solloway, M., O'Neill, J. W., Emery, J. & Bier, E. *Genes Dev.* **8**, 2602–2616 (1994).
- François, V. & Bier, E. *Cell* **80**, 19–20 (1995).
- Wieschaus, E., Nüsslein-Volhard, C. & Jürgens, G. *Wilhelm Roux Arch. dev. Biol.* **193**, 296–307 (1984).
- Zusman, S. B., Sweeton, D. & Wieschaus, E. F. *Dev. Biol.* **129**, 417–427 (1988).
- Ferguson, E. L. & Anderson, K. V. *Cell* **71**, 451–461 (1992).
- Wharton, K. A., Ray, R. P. & Gelbart, W. M. *Development* **117**, 807–822 (1993).
- Holley, S. A. et al. *Nature* **376**, 249–253 (1995).
- Fainsod, A., Steinbeisser, H. & De Robertis, E. M. *EMBO J.* **13**, 5015–5025 (1994).
- Dale, L., Howes, G., Price, B. M. & Smith, J. C. *Development* **115**, 573–585 (1992).
- Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. & Hogan, B. L. *Development* **115**, 639–647 (1992).
- Ferguson, E. L. & Anderson, K. V. *Development* **114**, 583–597 (1992).
- Sasai, Y., Lu, B., Steinbeisser, H. & De Robertis, E. M. *Nature* **376**, 333–336 (1995).
- Quiring, R., Walldorf, U., Kloter, U. & Gehring, W. J. *Science* **265**, 785–789 (1994).
- Halder, G., Callaerts, P. & Gehring, W. J. *Science* **267**, 1788–1792 (1995).
- Zuker, C. S., Cowman, A. F. & Rubin, G. M. *Cell* **40**, 851–858 (1985).
- Zuker, C. S. *Science* **265**, 742–743 (1994).
- Padgett, R. W., St Johnson, R. D. & Gilbert, W. M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2905–2909 (1993).
- Steinbeisser, H., Fainsod, A., Niehrs, C., Sasai, Y. & De Robertis, E. M. *EMBO J.* **14**, 5230–5243 (1995).
- Winnier, G., Blessing, M., Labowsky, P. A. & Hogan, B. L. M. *Genes Dev.* **9**, 2105–2116 (1995).
- Serafini, T. et al. *Cell* **78**, 409–424 (1994).

- Colamarino, S. A. & Tessier-Lavigne, M. *Cell* **81**, 621–629 (1995).
- Wilson, P. A. & Hemmati-Brivanlou, A. *Nature* **376**, 331–333 (1995).
- Nellen, D., Affolter, M. & Basler, K. *Cell* **78**, 225–237 (1994).
- Staebling-Hampton, K., Hoffmann, F. M., Baylies, M. K., Rushton, E. & Bate, M. *Nature* **372**, 783–786 (1994).
- Frasch, M. *Nature* **374**, 464–467 (1995).
- De Robertis, E. M. *Nature* **374**, 407–408 (1995).
- Smith, W. C., Knecht, A. K., Wu, M. & Harland, R. M. *Nature* **361**, 547–549 (1993).
- Lamb, T. M. et al. *Science* **262**, 713–718 (1993).
- Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. *Cell* **77**, 283–295 (1994).
- Grunz, H. & Tacke, L. *Cell Differ. Dev.* **28**, 211–217 (1989).
- Sato, S. M. & Sargent, R. D. *Dev. Biol.* **134**, 263–266 (1989).
- Haeckel, E. Q. *J. microsc. Sci.* **14**, 142–247 (1874).
- Slack, J. M. W., Holland, P. W. H. & Graham, C. F. *Nature* **361**, 490–494 (1993).
- Carroll, S. B. *Nature* **376**, 479–485 (1995).
- De Robertis, E. M. in *Guidebook of Homeobox Genes* (ed. Duboule, D.) 11–23 (IRL, Oxford, 1994).
- Scott, M. P. *Cell* **79**, 1121–1124 (1994).
- Phillipe, H., Chenail, A. & Adoutte, A. *Development* (suppl.) 15–25 (1994).
- Willmer, C. H. *Invertebrate Relationships* (Cambridge University Press, Cambridge, 1990).
- Preobrazhenskii, A. A. & Glinka, A. V. *Dokl. Akad. Nauk. SSSR* **284**, 1489–1491 (1985).
- Schmidt, J. E., Suzuki, A., Ueno, N. & Kimelman, D. *Dev. Biol.* **169**, 37–50 (1995).
- Hemmati-Brivanlou, A. & Thomsen, G. H. *Dev. Genet.* **17**, 78–89 (1995).
- Clement, J. H., Fettes, P., Knöchel, S., Lef, J. & Knöchel, W. *Mech. Dev.* **52**, 357–370 (1995).
- Xu, R.-H. et al. *Biochem. biophys. Res. Comm.* **212**, 212–219 (1995).
- Suzuki, A., Shioda, N. & Ueno, N. *Dev. Growth & Differ.* **35**, 581–588 (1995).
- Hawley, S. H. B. et al. *Genes Dev.* **9**, 2913–2935 (1995).
- Yamada et al. *J. Cell Biol.* **130**, 217–266 (1995).
- Schmidt, J., François, V., Bier, E. & Kimelman, D. *Development* **121**, 4319–4328 (1995).

ACKNOWLEDGEMENTS. The references cited in this Progress article are not comprehensive owing to space constraints. We thank L. Gont and L. Leys for comments on the manuscript. Our work is supported by a grant from the NIH. E.M.D.R. is an HHMI Investigator. Y.S. was an HFSP postdoctoral fellow and is an HHMI Research Associate.

The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor

David W. Banner*, Allan D'Arcy*, Christiane Chène*, Fritz K. Winkler*, Arabinda Guha†, William H. Konigsberg‡, Yale Nemerson†§ & Daniel Kirchhofer*

* Pharma Division, F. Hoffmann-La Roche, Grenzacherstrasse 124, 4002 Basle, Switzerland

Departments of † Medicine and § Biochemistry, Mount Sinai School of Medicine, New York, New York 10029, USA

‡ Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510, USA

Blood coagulation is initiated when tissue factor binds to coagulation factor VIIa to give an enzymatically active complex which then activates factors IX and X, leading to thrombin generation and clot formation. We have determined the crystal structure at 2.0-Å resolution of active-site-inhibited factor VIIa complexed with the cleaved extracellular domain of tissue factor. In the complex, factor VIIa adopts an extended conformation. This structure provides a basis for understanding many molecular aspects of the initiation of coagulation.

BLOOD coagulation is driven by the formation of complexes between serine protease coagulation factors and their corresponding membrane-bound cofactors. Coagulation factor VIIa (FVIIa) complexes with tissue factor (TF); factor IXa (FIXa) with factor VIIIa (FVIIIa); and factor Xa (FXa) with factor Va (FVa) (where 'a' denotes the product of enzymatic activation). The first of these enzyme-cofactor complexes to be crystallized is that of FVIIa and TF¹. FVII circulates as a single-chain zymogen of 406 residues consisting of an amino-terminal γ -carboxy glutamic acid-rich (Gla) domain, followed by two epidermal growth factor (EGF)-like domains, a short linker peptide, and a carboxy-terminal serine protease domain. The active enzyme form, FVIIa, is generated by specific cleavage of a peptide bond between Arg and Ile at the end of the linker peptide, giving an N-terminal light chain of 152 residues and a heavy chain of 254 residues, linked by a disulphide bridge. FVIIa alone shows very little proteolytic activity, and only realizes its full enzymatic potential when complexed to TF in the presence of Ca^{2+} (reviewed in refs 2–4). TF is located mainly in the adventitia of the vessel wall, and is also expressed in atherosclerotic plaques⁵. Vessel injury or plaque rupture exposes TF to FVII/FVIIa in the circulating blood, resulting in the formation of an active TF:FVIIa complex which initiates blood coagulation. FIX and FX are the substrates of the TF:FVIIa complex, have a high degree of sequence similarity to FVII, and the same overall architecture³. The structure of soluble tissue factor (sTF) (the extracellular part of tissue factor, residues 1–219) had already been determined^{6,7} before this study; sTF consists of two fibronectin type III-like domains, TF1 and TF2, which are arranged end to end. An extensive interface between the domains creates a rigid structure with an 'elbow' angle of about 120°. Published data on TF mutants had defined the individual amino acids essential for binding to FVIIa, showing that the main binding site for FVIIa is located at the interface between the domains TF1 and TF2 (refs 8–14). No structural data were available for any of the four domains of FVIIa, but models could be constructed from structures of related proteins^{15–19}. Several studies had shown that all four FVII domains are important for TF binding^{20–25}, and natural point mutations leading to dysfunction had been mapped to the EGF-1, EGF-2 and catalytic domains (reviewed in ref. 26). Even taken together, these data

were not sufficient to enable the construction of detailed models of FVIIa, of the TF:FVIIa complex, or of the mode of interaction of the complex with substrates.

Here we present the 2.0-Å X-ray crystal structure of the complex of active-site-inhibited FVIIa with subtilisin-treated sTF. The FVIIa molecule is seen to adopt a well-defined, extended conformation in the complex, with the structures of all four domains, including the Ca^{2+} binding sites, well determined. sTF maintains the same conformation as in the free form, and imposes on the otherwise flexible FVIIa molecule a specific overall conformation by binding to it in three distinct areas. The structure and spatial location of the active-site region provide insight into the way that TF turns FVIIa into a fully functioning enzyme. We conclude that the macromolecular substrates, FIX and FX, could adopt during catalysis an overall structure similar to that seen here for FVIIa in the enzyme complex.

Structure of the TF:FVIIa complex

Crystals of the TF:FVIIa complex were produced as previously described¹. Human recombinant FVIIa was preinhibited with D-Phe-L-Phe-Arg chloromethyl ketone (pFFRCMK). Human recombinant sTF was subjected to limited subtilisin digestion, generating two fragments, 5–84 and 90–210, which remain together despite the loss of residues 85–89. This cleaved sTF was found to facilitate crystallization of the complex without significant loss of catalytic activity¹, and will also be referred to as 'TF'. Synchrotron data were collected from a single crystal to 2.0-Å resolution (Table 1). The structure of the complex was solved by molecular replacement, using a thrombin-based model and coordinates of sTF. Our model is complete apart from the last ten residues of the FVIIa light chain and two loops of TF. It contains nine Ca^{2+} ions, and has been refined to a crystallographic *R*-factor of 23.2% at 2.0-Å resolution with good geometry (Table 1).

Two views of the complex are shown in Fig. 1, where the C terminus of TF, and thus the approximate position of the cell membrane, is at the bottom. The complex is about 115 Å long and has a diameter of 40–50 Å. This elongated shape was predicted from sedimentation and fluorescence-anisotropy studies²⁷. FVIIa adopts an extended conformation and wraps around TF with the

FIG. 1 Orthogonal views of the TF:FVIIa complex. The catalytic serine protease domain of FVIIa is shown at the top in blue with one disulphide loop highlighted in cyan (see text). The inhibitor in the active site is represented by ball-and-stick with green atoms. The light chain of FVIIa runs from the bottom to the top with the Gla domain cyan, the EGF-1 domain yellow, the short EGF-1–EGF-2 linker light green, and the EGF-2 domain and C-terminal peptide orange. Tissue factor N-terminal domain (TF1) is magenta, and tissue factor C-terminal domain (TF2) is green. Calcium atoms are red and disulphides yellow. The plane of the cell membrane is considered to be at the bottom of the picture. All figures were produced using the programs MOLSCRIPT⁴⁹, RAS-TER3D⁵⁰ and RENDER⁵¹.



Gla domain proximal to, and the catalytic domain distal to, the cell-membrane region. The extensive region of contact between FVIIa and TF can be divided into three main regions which together have a total buried surface area on complex formation of $1,800 \text{ \AA}^2$ (Table 2). The individual residues found in these contact interfaces correlate well with data from mutational analysis^{8–14}.

Tissue factor

The structure of TF in the TF:FVIIa complex is very similar to that of uncomplexed sTF^{6,7}, and is shown in Fig. 2. For the complexed and uncomplexed models, 170 α -carbons may be superimposed with a root-mean-square deviation (r.m.s.d.) of only $0.53 \text{ \AA}$. The N and C termini (<5 , >210) and the cleaved loop (85–89), missing here as a result of the subtilisin treatment, were also poorly defined in the uncomplexed structure^{6,7}. A small shift in the TF loop including Leu 133 and Phe 140 helps bring these two residues

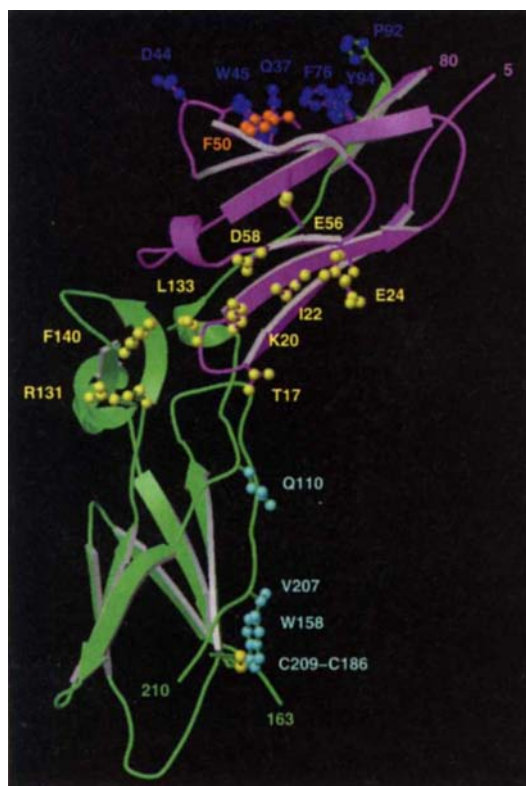


FIG. 2 TF as seen in the complex with FVIIa. The orientation is that of the left view in Fig. 1, and corresponds to that of ref. 7. The N-terminal domain (TF1) is on the upper right, and the C-terminal domain (TF2) at the bottom. Proteolytic cleavage in the loop between residues 80 and 92, used to facilitate crystal growth, has split the TF molecule into two fragments¹ (see text). The smaller, N-terminal fragment is shown magenta, and the larger, C-terminal fragment green. The peptide chain may be traced continuously from residues 5 to 80, 92 to 158, and 163 to 210 as shown. Side chains are shown only for those residues of TF which play a significant role in FVIIa binding, as determined from the structure of the complex. These residues may be grouped into three regions, the lower right-hand side of TF2, which binds the Gla domain of factor VIIa (Fig. 3a); the front central region at the TF1/TF2 domain boundary, which binds the EGF-1 domain (Fig. 3b); and the top of TF1, which binds the EGF-2 and catalytic domains (Fig. 3c). The colour scheme of Fig. 1 is used to indicate the FVIIa domain with which each residue interacts.

in contact with FVIIa. Residues 135–138 show some disorder, as do 158–164 and 181–183 at the bottom of the molecule. The side chains of TF that are important for binding of FVIIa, as determined from the structure of the TF:FVIIa complex, are shown in Fig. 2. The 'binding epitope' can be seen as a stripe running along the whole length of TF. The analysis below is in terms of three binding regions: the lower right-hand-side of TF2, which binds the Gla domain of FVIIa; the front central region at the TF1/TF2 domain boundary, which binds the EGF-1 domain; and the top of TF1, which binds the EGF-2 and catalytic domains. The correlation with the results of mutant work is remarkably good, given the very different methodologies used. Comparison of TF with the interferon (IFN)- γ receptor in complex with its ligand²⁸ shows the structures to be very similar, but the ligand-binding epitopes do not overlap; IFN- γ binds to loops on the receptor domain interface, but on the 'outside of the elbow', as does human growth hormone in the complex with its own receptor²⁹. Potential *N*-glycosylation sites on TF (Asn 11, Asn 124 and Asn 137) point away from FVIIa.

The light chain of FVIIa

The Gla domain of FVIIa is very similar to that of prothrombin fragment (PF) 2, in 2PF2.PDB¹⁷ (r.m.s.d. for 40 α -carbon pairs, 0.63 Å). Seven bound calcium ions are observed, six of which are arranged in a linear fashion (Fig. 1), and we can confirm γ -carboxylation for the first nine glutamic acid residues, with Glu 35 being disordered and possibly not modified, in accordance with biochemical data³⁰. Binding of Gla 6 and Gla 7 to the row of calcium ions makes a 'W-like' structure in which Phe 4, Leu 5 and Leu 8 protrude down towards the cell membrane. It remains to be shown whether these residues insert into the membrane, putting the row of six calcium ions in the plane of the phospholipid head-groups, as suggested for FIX and FX^{31,32}. In the TF:FVIIa complex, the Gla domain binds to TF2, as shown in Fig. 3a. The interaction is mainly hydrophobic, and involves mostly residues of the C-terminal helix of the Gla domain (the 'hydrophobic stack'²³) (Table 2).

The first EGF-like domain, EGF-1, binds a single calcium ion³³, as do the corresponding domains in FIX and FX^{34,35}. This is the first crystal structure of such a domain with calcium bound. The coordination sphere of the calcium ion is octahedral, having six oxygen ligands: Gln 49:OE1, apical; and Gln 64:O, Gly 47:O, Asp 46:OD2, Asp 63:OD1 and Asp 63:OD2, equatorial. These are the same as found by nuclear magnetic resonance for EGF-1 of FX¹⁹. One apical ligand is missing, and is presumably a water molecule. The nearest atom on either the Gla domain or TF is ~6 Å from the calcium. A direct functional role of the calcium ion in the EGF-1 domain is thus not apparent. One sugar, pointing away from TF, may be identified on each of the two known *O*-glycosylation sites in EGF-1, these being Ser 52 and Ser 60 (ref. 36). As shown in Fig. 3b, the EGF-1 domain contacts both TF domains and interacts with the TF residues most clearly identified by mutational analysis as being responsible for FVIIa affinity.

TABLE 1 Crystallographic data and structure solution

<i>(a) Data collection</i>				
Resolution	Observations	Unique	R_{merge}	$I/\sigma I$
1.95 Å	162,150	54,736	7.6%	10.6
<i>(b) Molecular replacement solution</i>				
Model: serine protease domain, 239 amino acids, 8–3.5 Å.				
	Best	Next best		
Rotation solution	10.2 σ	8.4 σ		
Translation solution	18.7 σ	9.5 σ		
<i>R</i> -factor	52.5%	54.6%		
After rigid body	22.2 σ	51.7%		
Model: tissue factor, 218 amino acids, 8–3.5 Å.				
	Best	Next best		
Rotation solution	10.5 σ	8.6 σ	(TF alone)	
Translation solution	31.4 σ	16.2 σ	(+ serine protease)	
	48.6%	52.8%		
After rigid body	37.7 σ	46.8%		
<i>(c) Refinement statistics</i>				
<i>R</i> -factor	$R = 23.2\%$ for 4,783 atoms in 590 residues, 20.0–2.0 Å.			
Geometry:	r.m.s. bond length error 0.10 Å, r.m.s. angle error 2.1°.			

Crystals of the TF:FVIIa complex, grown as described, have space group $P2_12_12_1$ with $a = 70.65$ Å, $b = 82.55$ Å, $c = 126.5$ Å (ref. 1). Data from one large crystal of dimensions $0.12 \times 0.12 \times 1.25$ mm were collected with a 30-cm marresearch image plate on beamline 9.6 at the Daresbury synchrotron with detector distance 300 mm, wavelength 0.89 Å, counting time 60 s, frame size 1° and beam cross-section ~0.2 mm. Initially, diffraction was to better than 1.95 Å, but after 25 frames it had deteriorated linearly to ~3.0 Å. The crystal was translated to put a fresh area in the beam and data collection restarted four times to give five batches covering 100° rotation angle. Data were processed using DENZO and SCALEPACK⁴⁵. It was found that both frame R_{merge} and apparent mosaic spread increased linearly with exposure. Ultimately, 76 frames were accepted, each with resolution better than 2.5 Å, frame $R_{\text{merge}} < 12\%$ and $I/\sigma I > 8.0$. Statistics of the merged data are given above. The data are >95% complete in all resolution shells. In the outer shell $I/\sigma I$ is ~2.0 and a resolution cut-off of 2.0 Å was applied to give 49,724 reflections. The structure was solved by molecular replacement using the program AMORE⁴⁶ in the CCP4 program package⁴⁷. A thrombin-based model of the serine protease domain and the sTF model obtained from Y. Muller and A. de Vos⁷ could be correctly oriented using the rotation function and then brought on to the same origin with the translation function (statistics given above). Positional and restrained *B*-value refinement was performed with X-PLOR⁴⁸ and the missing domains located and built using the in-house graphics program MOLOC of P. Gerber. The current refinement statistics are given above. All residues fall in allowed regions of the Ramachandran plot. The present model has 590 amino acid residues (of which 9 are Gla and 3 the inhibitor), 9 Ca^{2+} ions, 1 cacodylate and 198 water molecules. Residual density is seen for one sugar on each of Ser 52L and Ser 60L and for Glu 91T. Some improvement of side-chain conformations and water structure will be carried out. All figures were produced using the programs MOLSCRIPT⁴⁹, RASTER3D⁵⁰ and RENDER⁵¹.

The second EGF-like domain, EGF-2, has the same structure as seen for EGF-2 in the recently released FX structure, 1HCG.PDB³⁷ (r.m.s.d. for 47 α -carbon pairs, 0.92 Å), and interacts with the catalytic domain in the same way. The EGF-2 and catalytic domains together form a coherent structural unit, and interact with TF1 as shown in Fig. 3c. A short, poorly formed helix connects the two EGF-like domains. This could act as a hinge region in uncomplexed FVIIa, as could the connection between the EGF-1 and Gla domains, consistent with the observed segmental domain motion of free FVIIa which is frozen in the TF:FVIIa complex²⁷. The lack of an interface between the EGF domains contrasts with the TNF receptor³⁸, where the domains interact strongly end-to-end to give a linear structure.

The catalytic domain

The catalytic domain (Fig. 4) has a core structure common to all serine proteases of the thrombin/trypsin family (r.m.s.d. for 200 α -carbon pairs, 0.81 Å to D-Phe-L-Pro-Arg chloromethyl ketone (PPACK)-thrombin^{15,16}) and the residues involved in primary

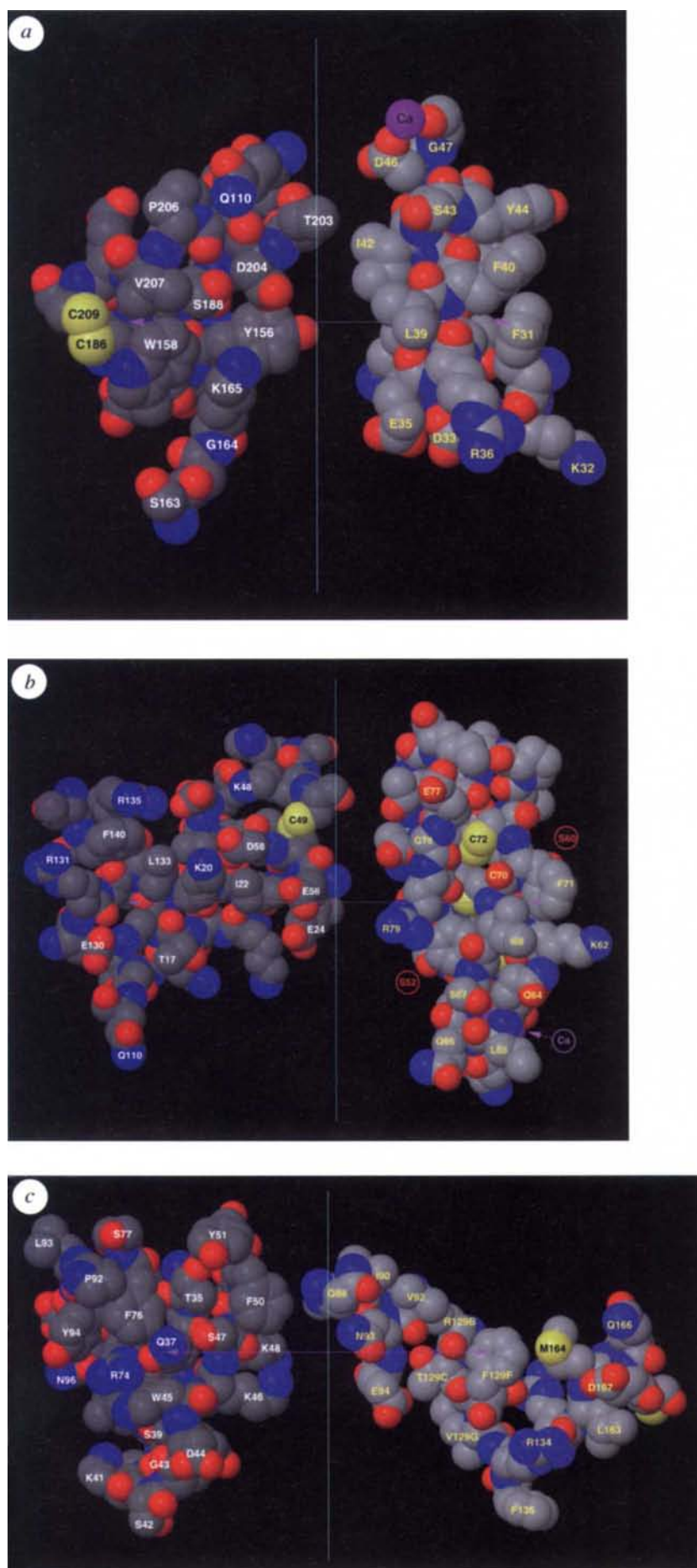


FIG. 3 Space-filling representations of the contact surfaces of TF (left, darker) with FVIIa (right, lighter). In each case the surfaces have been opened out like a book about the vertical line, one major interaction being indicated by the magenta arrow. To differentiate between the residues in the different molecules, the chain identifier follows each residue number below, with T for tissue factor, L for FVIIa light chain, and H for FVIIa heavy chain. The standard serine protease numbering¹⁵ is given in curly brackets, and may have an insertion code; for example, Phe 275H {129F} is the 275th residue of FVIIa and the 6th residue of an insertion loop following residue 129 in trypsin/chymotrypsin. **a**, Contact surface between TF C-terminal domain (left) and FVIIa Gla domain (right). The 'hydrophobic stack' helix of the Gla domain runs vertically from D33L towards G47L. The first backbone hydrogen bond of the helix is from Lys 38L:NH to Ala 34L:O, with Asp 33L 'capping' the helix N terminus with hydrogen bonds Arg 36L:NH to Asp 33L:OE1 and Thr 37L:OG to Asp 33L:O. The last hydrogen bond in the current model is from Asp 46L:NH to Ser 43L:O (a typical 3:10 C terminus), but there is a certain amount of disorder in this region, and it is possible that an Asp 46L:N to Ile 42L:O hydrogen bond also exists some of the time. The Ca^{2+} ion (magenta) from the EGF-1 domain is also shown with the first two of its contact residues, Asp 46L and Gly 47L. The exact boundary between the Gla and EGF-1 domains is a matter of definition, but the connection is weak, as it consists only of the hydrogen bonding from Asp 46L:NH to the C terminus of the hydrophobic stack. The major hydrophobic interaction between TF and FVIIa is Val 207T, Trp 158T and the Cys 209T to Cys 186T disulphide bridge with Phe 31L (magenta arrow). Val 207T also contacts Phe 40L. The rest of the Gla domain would be to the right of the K32–Y44 line, and does not contact TF, except that the Cys 209T–Cys 186T disulphide bridge may be further shielded from solvent by the side chain of L13L, which contacts Phe 31L and Phe 41L, but is poorly ordered and not shown here. The only well-determined interface hydrogen bond is Q110T:NE2 to S43L:OG. Residues Ser 163T, Gly 164T, Lys 165T, Trp 158T and Arg 36L, Leu 39L, Glu 35L (possibly Gla) are all poorly ordered, so it is difficult to say whether the interaction of Arg 36L with TF is significant. **b**, Contact surface between the TF domain interface region (left) and FVIIa EGF-1 domain (right). TF is oriented as in Fig. 2, so TF1 is upper right and TF2 lower left. The major hydrophobic interactions of TF with FVIIa are: Phe 140T, Leu 133T and Arg 131T with Phe 71L (magenta arrow); Leu 133T, Lys 20T and Thr 17T with Ile 69L; and Ile 22T with Arg 79L. Lys 20T points into the hole just below the Cys 72L to Cys 81L disulphide bridge. Well-determined hydrogen bonds involving TF1 are: Lys 20T:NZ to Cys 70L:O, Lys 20T:NZ to Gly 78L:O, Lys 48T:NZ to Glu 77L:OE2, Asp 58T:OD2 to Gly 78L:NH, Glu 56T:OE1 to Arg 79L:NE, Glu 56T:OE2 to Arg 79L:NH1, Glu 24T:OE1 to Arg 79L:NH1. The only hydrogen bond from TF2 is Gln 110T:NH to Gln 64L:OE1 (Arg 135T and Glu 130T have weak electron density). Residues Arg 79L and Lys 20T are important as they participate in both hydrophobic and hydrogen bonding interactions. The Ca^{2+} ion (magenta) is depicted with increased radius to make it visible in this view. There is clearly no direct interaction with TF, but there is an indirect one through Gln 64L. O-glycosylated residues Ser 52L and Ser 60L are marked. The carbohydrate on Ser 52L may interact with residues providing Ca^{2+} ligands, and there is some disorder in this area. **c**, Contact surface between the TF N-terminal domain (left) and FVIIa EGF-2 and catalytic domains (right). TF is viewed down from the top relative to Figs 1 and 2 and the β -strands are close to vertical here (for example, Thr 35T to Ser 39T). Residues Gln 88L to Glu 94L, which

form the bottom end of the EGF-2 domain of FVIIa, may be seen near the centre of the figure. To their right are 8 residues Arg 271H{129B} to Phe 278H{135}, the first 6 of which form the C terminus of an α -helix. At the far right are residues Arg 304H{162} to Cys 310H{168}, the last 4 of which form the N terminus of another α -helix (see Fig. 1, right view). This is rather a complex interface region, in which five key interactions may be identified: (1) the side chain of Phe 50T is buried by residues Gln 88L to Asn 93L; (2) residues Phe 76T, Tyr 94T, Pro 92T and Arg 74T form a hydrophobic pocket which binds the side chain of Met 306H{164}; (3) Tyr 94T interacts with Gln 308H{166} and Asp 309H{167} by 3 hydrogen bonds; (4) Trp 45T packs against the side chain of Arg 277H{134}, which is further stabilized by two hydrogen bonds to Ser 39T and Gly 43T; and (5) Gln 37T participates in an unusual interaction in that its side chain is held fixed by hydrogen bonding with Arg 74T and Ser 47T, and a hydrogen bond is donated to the pi electrons of the phenyl ring of Phe 275H{129F} (magenta arrow). There are many water molecules in this region; for example, three accept hydrogen bonds from the side chains of Arg 74T, three more from Arg 304H{162}, and one from Trp 45T. The intermolecular hydrogen bonds are Tyr 94T:O to Gln 308H:NE2{166}, Tyr 94T:OH to Asp 309N:NH{167}, Tyr 94T:OH to Asp 309H:OD2{167}, Asn 96T:ND2 to Asp 309H:OD1{167}, Ser 39T:OG to Arg 277H:NH2{134}, Gly 43T:O to Arg 277H:NH2{134}, and Trp 45T:NH to Phe 275H:O{129F}. This last hydrogen bond should be noted as the only main chain-main chain contact.

substrate recognition and in catalysis are highly conserved both in sequence and spatial relationship. The inhibitor chosen for crystallization studies, α -FFRCMK, binds to the active-site histidine and serine exactly as the analogous PPACK binds to thrombin. A high-affinity calcium-binding site is found in the FVIIa catalytic domain, as predicted^{33,39}. This is similar to that found in trypsin, except that Glu 77 of trypsin is here replaced by Asp, which binds to the calcium through a water molecule. All of the other surface loops, however, differ from previously determined structures. Of primary relevance to substrate recognition is the sequence Pro-Gly-Thr-Thr, 236-239, {96-99} (here and elsewhere the chymotrypsinogen-based numbering scheme is given in curly brackets, see ref. 15). This loop defines the nature and extent of the hydrophobic substrate/inhibitor binding pockets (S2 and S3 subsites, P- and D- pockets¹⁶), which are clearly different from FIX, FX and thrombin, which all have different substrate specificities.

Activation

The end of the EGF-2 domain and the ends of two short helical stretches of the catalytic domain together form an extended concave surface which arches over the whole top of the TF1 domain (Figs 1 and 3c). One of these sequences, Met-Thr-Gln-Asp, 306-309, {164-167} precedes the loop, Cys 310-Cys 329 {168-182}, highlighted in Figs 1 and 4. This loop is five residues longer than in the related serine proteases, and binds a cacodylate ion from the crystallization buffer. This is the only potentially mobile structure between the FVIIa active site and TF, and is therefore a prime candidate for involvement in the activation process. By using the doubly covalent inhibitor we have forced the active-site region into an 'active' conformation. We do not know, however, the precise relationship of the structure observed here to that of free FVIIa or to the conformation in the presence of macromolecular substrates. Modelling shows that the 168-182 loop is just long enough to be able either to interfere with substrate binding in the region of the D-Phe present here, or to interact with the N-terminal Ile 153 {16}, two possible modes of autoinhibition (a third being that, in the absence of TF, the EGF-1 and/or Gla domains reduce substrate access to the active site).

A detailed inspection of this interface region (Fig. 3c, legend) reveals only two small, separated, hydrophobic 'lock-and-key' interactions (the side chain of Phe 50 of TF is trapped in a pocket formed by the end of the EGF-2 domain, and the side chain of Met 306 {164} is enclosed by TF residues 74, 76, 92 and

94. Otherwise, a variety of hydrogen-bonding interactions are seen, involving main-chain and side-chain atoms, and also the 11 water molecules so far identified. In particular, the centre of this region contains water, which contrasts with the other two interface regions, each of which has a central hydrophobic patch similar in nature to that found in the growth hormone/growth hormone receptor complex²⁹ (see ref. 40 for detailed analysis). This FVIIa catalytic domain/TF interface does not merely contribute to affinity, but has the function of mediating enzyme activation, a complex process probably involving a number of substantial conformational changes. It is possible that hydrophilic interactions are more suitable than hydrophobic ones to allow this functionality.

Interaction with substrates

In the orientation of the complex shown in Fig. 1, the inhibitor (and hence also a peptidic substrate) runs horizontally about 80 Å above the row of calcium ions in the Gla domain, which we take to lie approximately in the plane of the cell membrane. The similarity of the domain structure of FIX and FX to that of FVII caused us to examine the hypothesis that FIX and FX, as substrates, adopt during their proteolytic activation an extended conformation similar to that of FVIIa seen here. If a model of FVII is constructed from FVIIa in the orientation shown, then both the scissile bond (Arg 152-Ile 153) {16} and the last observed residue of the light chain, Glu 142, lie about 70 Å from the cell membrane. Although both enzyme (TF:FVIIa) and substrate (FIX, FX) are expected to be highly mobile, both in the plane of the membrane and relative to it, there is an absolute structural requirement that the substrate runs correctly through the active

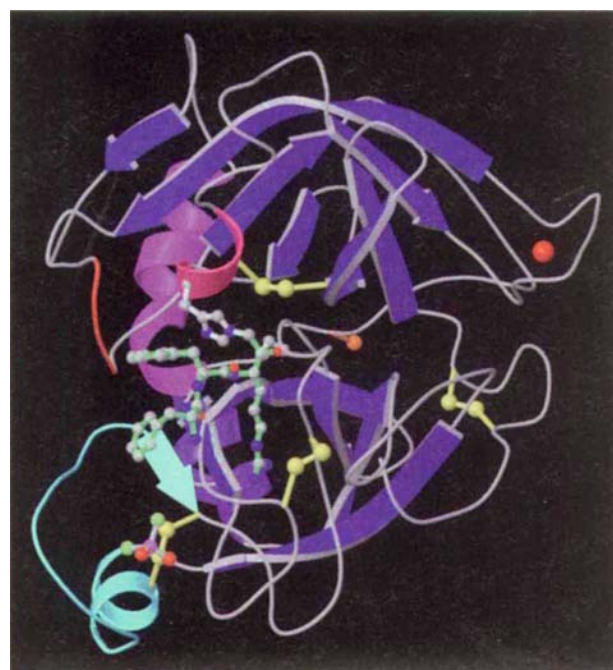


FIG. 4 Ribbon diagram of the catalytic domain of FVIIa, with β -sheets shown in blue, α -helices in indigo for 266-275{126-129F}, magenta for 380-391{231-242}, and crimson for 192-195{56-59}, disulphides in yellow, the Ca^{2+} ion in red, the inhibitor as ball-and-stick with green bonds, and the active-site histidine and serine side chains as ball-and-stick with white bonds. Residues Thr 307{165} to Cys 329{182}, possibly involved in activation, are shown in cyan with the bound cacodylate ion as ball-and-stick. The upper left β -strand is a continuation of the 'C-terminal' helix. Below it in red is the loop 236-239 {96-99} with sequence Pro-Gly-Thr-Thr, which builds the left side of the inhibitor (substrate) binding pockets and confers specificity for short peptidic substrates. The attachment to the light chain is by Cys 262{122} shown in orange (rear, centre). The N terminus is at the lower right between two disulphides.

site during catalysis. This can be achieved very easily for the above models by tilting about 20–30° around axes running horizontally through the calcium region of the Gla domain, plus the appropriate lateral rearrangement (a direct geometrical consequence of the inhibitor and the peptide to be cleaved (up to Arg 152) both lying roughly along the 'equator' of the catalytic domain and running round it in the opposite sense). By replacing the modelled FVII by FIX or FX, a broad family of substrate models is obtained, some of which have reasonable contact between the Gla domain of the modelled substrate and the membrane region of the TF:FVIIa enzyme complex. This modelling exercise suggests that FVII, and possibly also TF:FVII, could act as substrates for TF:FVIIa. Whether FVII alone or in complex with TF is a substrate of TF:FVIIa is unclear (for a discussion see ref. 4). Although plausible models of FIX and FX as substrate can be built, a difference in sequence length between FVII, FIX and FX occurs in the short peptide linking the EGF-1 and EGF-2 domains. A kink was found here in the recent structure determination of calcium-free FIX⁴¹, and it remains to be seen if this persists in enzyme/substrate complexes. The exact orientation of the molecules with respect to the membrane surface cannot be determined from the crystal structure, but an arrangement more or less perpendicular to the membrane would be expected from fluorescence studies in which the distances of FXa and FIXa active sites from the membrane surface are reported to be at least 60 and 70 Å, respectively^{42,43}.

The dramatic increase in the catalytic efficiency of FVIIa for activation of FIX or FX that occurs in the presence of TF⁴⁴, coupled with high substrate specificity, may now be appreciated as the result of a multistep process. FVIIa is first captured onto the membrane surface by TF, thereby losing much of its rotational and translational freedom as well as its internal flexibility. The membrane-bound substrate is then captured by the TF:FVIIa complex. This must bring the peptide to be cleaved correctly into the active site, which may already be in its optimum conformation for catalysis owing to the interaction with TF. Alternatively, there may be a further conformational change on substrate binding, perhaps also mediated by TF. Only when the ternary complex has been properly assembled can catalysis finally occur.

TABLE 2 Structural data of the FVIIa:TF complex

Size: length, 115 Å; diameter, 40–50 Å	
Intermolecular chain contact areas (Å ²)	
Tissue factor–factor VIIa	1,810
Tissue factor–light chain	1,340
Tissue factor–heavy chain	470
Intermolecular domain contact areas (Å ²)	
TF1–EGF-1	440
TF1–EGF-2	203
TF1–serine protease	475
TF2–Gla	410
TF2–EGF-1	383
Intramolecular domain contact areas (Å ²)	
Heavy chain–light chain	990
TF1–TF2	562
Definitions	
TF1, tissue factor 5–106	
TF2, tissue factor 107–210	
Gla, factor VII 1–45	
EGF-1, factor VII 46–86	
EGF-2, factor VII 87–142	

Contact surface areas quoted are the change in solvent-accessible surface area on complex formation, computed by the program XSAE of C. Broger, with a probe radius of 1.4 Å. The area changes on the two interacting surfaces are not identical owing to edge effects. Here the average of the two is given (not the sum).

The structure presented here provides a basis for understanding many aspects of the initiation of coagulation, and should act as a stimulus to a wide range of experiments. We anticipate that it will also facilitate the design of TF:FVIIa inhibitors as novel anti-coagulants for the treatment of a wide range of thrombotic diseases. □

Received 20 November 1995; accepted 6 February 1996.

- Kirchhofer, D. et al. *Proteins* **22**, 419–425 (1995).
- Nemerson, Y. *Blood* **71**, 1–8 (1988).
- Davie, E. W. *Thromb. Haemostasis* **74**, 1–6 (1995).
- Rapaport, S. I. & Rao, L. V. M. *Thromb. Haemostasis* **74**, 7–17 (1995).
- Wilcox, J. N., Smith, K. M., Schwartz, S. M. & Gordon, D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2839–2843 (1989).
- Harios, K. et al. *Nature* **370**, 662–666 (1994).
- Muller, Y. A., Ultsch, M. H., Kelley, R. F. & de Vos, A. M. *Biochemistry* **33**, 10864–10870 (1994).
- Ruf, W., Schullek, J. R., Stone, M. J. & Edgington, T. S. *Biochemistry* **33**, 1565–1572 (1994).
- Ruf, W. et al. *Biochemistry* **34**, 6310–6315 (1995).
- Schullek, J. R., Ruf, W. & Edgington, T. S. *J. biol. Chem.* **269**, 19399–19403 (1994).
- Gibbs, C. S., McCurdy, S. N., Leung, L. L. K. & Paborsky, L. R. *Biochemistry* **33**, 14003–14010 (1994).
- Kelley, R. F., Costas, K. E., O'Connell, M. P. & Lazarus, R. A. *Biochemistry* **34**, 10383–10392 (1995).
- O'Brien, D. P. et al. *Biochemistry* **33**, 14162–14169 (1994).
- Martin, D. M. A., Boys, C. W. G. & Ruf, W. *FASEB J.* **9**, 852–859 (1995).
- Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R. & Hofsteenge, J. *EMBO J.* **8**, 3467–3475 (1989).
- Banner, D. W. & Hadvary, P. J. *biol. Chem.* **266**, 20085–20093 (1991).
- Soriano-Garcia, M., Padmanabhan, K., de Vos, A. M. & Tulinsky, A. *Biochemistry* **31**, 2554–2566 (1992).
- Graves, B. J. et al. *Nature* **367**, 532–538 (1994).
- Selander-Sunnerhagen, M. et al. *J. biol. Chem.* **267**, 19642–19649 (1992).
- Toomey, J. R., Smith, K. J. & Stafford, D. W. *J. biol. Chem.* **266**, 19198–19202 (1991).
- Clarke, B. J. et al. *FEBS Lett.* **298**, 206–210 (1992).
- Sakai, T., Lund-Hansen, T., Thim, L. & Kisiel, W. *J. biol. Chem.* **265**, 1890–1894 (1990).
- Petersen, L. C., Schiødt, J. & Christensen, U. *FEBS Lett.* **347**, 73–79 (1994).
- Wildgoose, P., Kazim, A. L. & Kisiel, W. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7290–7294 (1990).
- Kazama, Y., Pastuszyn, A., Wildgoose, P., Hamamoto, T. & Kisiel, W. *J. biol. Chem.* **268**, 16231–16240 (1993).
- Tuddenham, E. G. D., Pemberton, S. & Cooper, D. N. *Thromb. Haemostasis* **74**, 313–321 (1995).

- Waxman, E., Laws, W. R., Laue, T. M., Nemerson, Y. & Ross, J. B. A. *Biochemistry* **32**, 3005–3012 (1993).
- Walter, M. R. et al. *Nature* **376**, 230–235 (1995).
- de Vos, A. M., Ultsch, M. & Kossiakoff, A. A. *Science* **255**, 306–312 (1992).
- Thim, L. et al. *Biochemistry* **27**, 7785–7793 (1988).
- Freedman, S. J., Fure, B. C., Fure, B. & Baleja, J. D. *Biochemistry* **34**, 12126–12137 (1995).
- Sunnerhagen, M. et al. *Nature struct. Biol.* **2**, 504–509 (1995).
- Schiødt, J., Harrit, N., Christensen, U. & Petersen, L. C. *FEBS Lett.* **306**, 265–268 (1992).
- Handford, P. A. et al. *EMBO J.* **9**, 475–480 (1990).
- Persson, E. et al. *J. biol. Chem.* **264**, 16897–16904 (1989).
- Bjoern, S. et al. *J. biol. Chem.* **266**, 11051–11057 (1991).
- Padmanabhan, K. et al. *J. molec. Biol.* **232**, 947–966 (1993).
- Banner, D. W. et al. *Cell* **73**, 431–445 (1993).
- Sabharwal, A. K. et al. *J. biol. Chem.* **270**, 15523–15530 (1995).
- Clackson, T. & Wells, J. A. *Science* **267**, 383–386 (1995).
- Brandstetter, H., Bauer, M., Huber, R., Lollar, P. & Bode, W. *Proc. natn. Acad. Sci. U.S.A.* **92**, 9796–9800 (1995).
- Husten, E. J., Esmon, C. T. & Johnson, A. E. *J. biol. Chem.* **262**, 12953–12961 (1987).
- Mutucumarana, V. P., Duffy, E. J., Lollar, P. & Johnson, A. E. *J. biol. Chem.* **267**, 17012–17021 (1992).
- Komiyama, Y., Pedersen, A. H. & Kisiel, W. *Biochemistry* **29**, 9418–9425 (1990).
- Otinowski, Z. In *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 56–62 (SERC Daresbury Laboratory, Warrington, 1993).
- Navaza, J. *Acta crystallogr.* **A50**, 157–163 (1994).
- Collaborative Computational Project, No. 4 *Acta crystallogr.* **D50**, 760–763 (1994).
- Brünger, A. T. *X-PLOR, Version 3.1. A System for X-Ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Bacon, D. J. & Anderson, W. F. *J. molec. Graph.* **6**, 219–220 (1988).
- Merritt, E. A. & Murphy, M. E. P. *Acta crystallogr.* **D50**, 869–873 (1994).

ACKNOWLEDGEMENTS. We thank Y. Muller and A. de Vos for TF coordinates; P. Riskalia for help in data collection; and P. Gerber, C. Broger and D. Bur for software support. Y.N., W.H.K. and A.G. acknowledge support from the United States Public Health Service. Coordinates will be deposited in the Brookhaven Protein Data Bank for release after one year.

Measurement of the opacity of ionized helium in the intergalactic medium

Arthur F. Davidsen, Gerard A. Kriss & Wei Zheng

Center for Astrophysical Sciences, Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA

DETERMINATION of the density and ionization state of diffuse primordial gas in the intergalactic medium (IGM) is an important step towards understanding the nature of dark matter and the formation of structure in the Universe. Absorption by neutral hydrogen atoms in the IGM should produce a large flux deficit in the spectra of high-redshift quasars at wavelengths below that of the H I Lyman- α line (the Gunn-Peterson effect¹). But observations of quasars at low spectral resolution have revealed only relatively small deficits, and at higher spectral resolution these appear as numerous discrete absorption features (the Lyman- α forest). This implies that intergalactic hydrogen is both highly ionized and non-uniformly distributed². Similar absorption is expected from intergalactic helium³, although its detection is more challenging and evidence for its existence was obtained only recently^{4,5}. Here we report the observation of absorption from singly ionized helium (He II) in the spectrum of the quasar HS1700 + 64. We measure a mean flux deficit over the redshift range $2.2 < z < 2.6$ that yields an effective optical depth of 1.00 (± 0.07). Comparing this result with the H I absorption towards this quasar, we argue that most of the measured He II absorption occurs in the lowest-density regions of the IGM.

Whereas hydrogen absorption at $z > 2$ can be observed with ground-based optical telescopes, He II Lyman- α has a rest wavelength of 304 Å and can only be observed with an ultraviolet telescope. Singly ionized helium gas is accessible only at $z > 2$, as the interstellar medium in our own Galaxy is opaque at wavelengths less than 912 Å. For $z > 3$ the Hubble Space Telescope (HST) may be used, but most such high-redshift quasars are undetectable at far-ultraviolet wavelengths⁶, primarily because there are normally several intervening absorbers with hydrogen that is optically thick in the Lyman continuum. Nevertheless, using the newly repaired HST and the objective prism mode of the Faint Object Camera, Jakobsen *et al.*⁴ succeeded in finding a quasar (Q0302 - 003, V-band magnitude $V = 17.8$, $z = 3.286$) with detectable flux down to $\sim 1,300$ Å (corresponding to a rest wavelength of ~ 304 Å), and no detectable flux in a 50-Å band below this wavelength. This was the first evidence of He II absorption in the IGM. The absence of any detected flux below the absorption edge produced by He II Lyman- α at the quasar redshift indicated that the effective optical depth $\tau_{\text{He II}} > 1.7$ at 90% confidence. Tytler *et al.*⁵ used the Faint Object Spectrograph (FOS) on the HST to obtain a spectrum of Q1935 - 69 ($V = 18.8$, $z = 3.170$), and reported He II absorption with the significantly smaller value $\tau_{\text{He II}} = 1.0$ (± 0.2), but this result could be affected by scattered light in the spectrograph, which produces a large background signal at the wavelengths of interest.

He II 304 Å absorption can be observed over the broad redshift range $2 < z < 3$ with a telescope and spectrograph specially designed to be sensitive to wavelengths in the 900–1,200 Å band. To carry out such measurements, we developed⁷ the Hopkins Ultraviolet Telescope (HUT) as a payload for the NASA/ESA Spacelab programme for the Space Shuttle. An initial attempt to observe He II absorption towards HS1700 + 64 was made with HUT on the Astro-1 mission⁸ in 1990, but pointing problems prevented acquisition of a spectrum with a sufficiently high signal-to-noise ratio. However, during the March 1995

reflight of HUT aboard the Astro-2 mission these observations were successfully completed. We selected the quasar HS1700 + 64 ($z = 2.72$ (but see below), $V = 16.1$) for observations with HUT because it is the brightest known quasar⁹ at observed wavelength 1,300 Å with $z > 2$. It is one of the rare high-redshift quasars that has no optically thick Lyman limit absorber anywhere in its spectrum, and its far-ultraviolet flux actually rises toward shorter wavelengths¹⁰. It is a factor of 10 brighter at far-ultraviolet wavelengths than the higher-redshift quasar Q0302 - 003 observed with HST⁴.

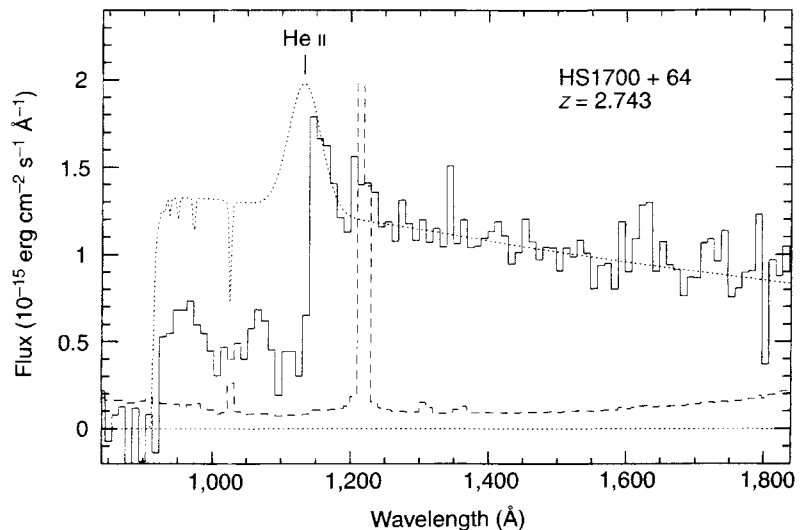
We observed HS1700 + 64 during the night-time portions of 12 orbits distributed throughout the Astro-2 mission, obtaining a total exposure of 19,507 s on the source through a 20-arcsec aperture. Pointing stability was typically good to 1 arcsec, as verified by HUT acquisition TV camera images of three guide stars in the field. The HUT spectrograph⁷ covers the wavelength range 820–1,840 Å at a resolution of ~ 3 Å. It employs a photon-counting microchannel-plate detector with 2,048 pixels, at a dispersion of 0.51 Å per pixel. The source provided a count-rate of ~ 1.5 counts s^{-1} , compared with a background of ~ 19.5 counts s^{-1} , concentrated mainly in the geocoronal Lyman- α emission line. We measured the dark count-rate (0.8 counts s^{-1}) on several occasions throughout the mission, and we measured the airglow line profiles (including their scattering wings, which include a broad-band background comparable to the dark rate) during several observations with no source in the aperture. These were then scaled and subtracted from the summed spectra of HS1700 + 64 to provide a 'cleaned count spectrum' of the source. An excellent check on the background subtraction is provided by the data in the region from 840 to 900 Å, where the source counts should be zero, as the interstellar medium is completely opaque at these wavelengths. We found the net count-rate in this band to be 6.6×10^{-5} counts s^{-1} ($\pm 1.9 \times 10^{-3}$ counts s^{-1}).

As an independent verification of the reliability of our background subtraction techniques, we subdivided our six blank sky observations into two sets. One set, totalling 4,008 s of integration through a $20'' \times 197''$ aperture, was used to determine the Lyman- α scattering profile. The remainder, totalling 6,780 s of integration through the same 20-arcsec-diameter aperture used to observe HS1700 + 64, was processed in the same way as the quasar data. This 'cleaned count spectrum' of the background has a net count rate in the 1,050–1,130 Å band of $(8.0 \pm 5.0) \times 10^{-3}$ counts s^{-1} compared to 4.1×10^{-2} counts s^{-1} for HS1700 + 64 in the same band.

The cleaned count spectrum of HS1700 + 64 was converted to absolute flux using the instrumental calibration¹¹ established by comparing observations of several DA white dwarfs with model stellar atmosphere calculations. The wavelength scale (accurate to ± 0.5 Å) and resolution (varying smoothly with wavelength from 2.0 to 4.5 Å) were established by observations of several symbiotic stars with very narrow emission lines. The resulting spectrum, shown in Fig. 1, clearly reveals absorption by intergalactic helium along the line of sight to HS1700 + 64. The quasar's power-law continuum emission is seen at wavelengths longer than 1,200 Å. Below this wavelength, the intensity first rises owing to redshifted He II 304 Å emission at the quasar redshift (with a possible contribution from O III 305 Å), and then falls abruptly between observed wavelengths 1,140–1,130 Å. At shorter wavelengths the depressed flux remains clearly detected down to the Galactic Lyman limit.

The usually quoted redshift of HS1700 + 64 is 2.72, but a re-examination of the optical data (performed independently and without knowledge of the results reported here) gives an emission line redshift of 2.730 (± 0.001) and a systemic redshift of $z_0 = 2.743$ (± 0.012), which includes an estimated offset of 1,000 ($\pm 1,000$) km s^{-1} in the quasar frame for the high-ionization emission lines (B. Espey, personal communication). This would

FIG. 1 Far-ultraviolet spectrum of quasar HS1700 + 64 observed with the Hopkins Ultraviolet Telescope, shown as a histogram with 10-Å bins. The dotted curve shows an extrapolation of the power-law spectrum ($F_\nu \propto \nu^{-1.0}$, where F_ν is the flux density at frequency ν) observed at wavelengths $>1,250$ Å, with an added emission feature for the He II 304 Å resonance line modelled as a gaussian (full width at half maximum, $14,500 \text{ km s}^{-1}$) at $z_e = 2.730$. Galactic reddening^{26,27} with $E_{B-V} = 0.02$ and interstellar hydrogen²⁸ Lyman series absorption for $N_H = 2 \times 10^{20} \text{ cm}^{-2}$ and $b = 10 \text{ km s}^{-1}$ have also been applied to the model. The extinction-corrected intensity of the He II 304 Å emission line in our spectrum is $5.4 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$. This is 8.3% of the intensity of Ly α as observed in the MMT spectrum, a ratio consistent with the range of 0.07–0.23 predicted in models of broad-line clouds in quasars^{29,30}. The dashed line shows the statistical error at each wavelength bin. The horizontal dotted line gives a zero flux reference across the spectrum. The Gunn–Peterson absorption trough arising from intergalactic He II and the He II Lyman- α forest is evident, extending from the systemic redshift ($z = 2.743$) to the Galactic Lyman limit at $z = 2$.



place He II 303.78 Å at an expected (systemic) wavelength of $1,137.0 (\pm 3.6)$ Å. We determined the wavelength of the observed absorption edge by assuming that it is intrinsically sharp, broadened only by the instrumental profile, employing the data at full resolution. We find the edge is at $1,138.2 (+0.6, -0.8)$ Å, corresponding to a redshift determined from the He II absorption $z_{\text{He II}} = 2.747 (\pm 0.003)$. This differs from the systemic redshift by $\Delta z = 0.004$, or $\sim 300 \text{ km s}^{-1}$ in the quasar frame. There is thus no difference, within the errors, between the best estimate of the systemic redshift of HS1700 + 64 and the redshift of the observed He II trough edge.

We also considered the possibility that the absorption edge is broadened by the proximity effect^{12,13}, but the observed edge is consistent with no broadening. An upper limit on the effective width of the proximity profile is 4.5 Å, or $\Delta z_p < 0.015$. The sharpness of the edge implies either that the IGM density in the vicinity of the quasar is very high, or that HS1700 + 64 is not producing much extra helium ionization in the IGM, at least along our line of sight¹².

The precision with which the redshift of the He II absorption feature matches the quasar's systemic redshift makes it very unlikely that it is produced by a low-redshift H I Lyman limit system that intervenes accidentally at the required redshift, $z = 0.247 (\pm 0.004)$. This same argument¹ was also used to make the case that He II was responsible for the absorption seen in Q0302 – 003. The case is even stronger here, because of the higher resolution and better wavelength accuracy of our result. Nevertheless, we have checked for other absorption lines that would be expected if the trough were due to a hydrogen Lyman limit system. We used the HUT data, HST/FOS data¹⁰, and an optical spectrum (R. Green, personal communication) taken with the Multiple Mirror Telescope (MMT) to search for Lyman- α and Lyman- β and the C IV and Mg II doublets, but no reasonable match was found. We conclude that the absorption we observe is due to ionized helium distributed in the IGM over redshifts from 2.0 to 2.7.

Low-resolution optical observations of H I absorption due to Lyman- α forest clouds and the Gunn–Peterson effect are usually characterized¹⁴ by the quantity D_A , the average fractional depression of the extrapolated quasar continuum in a wavelength region between Lyman- α and Lyman- β . To avoid the wings of the emission lines as well as the region near the quasar where the proximity effect may be important, it is customary to use¹⁵ the rest wavelengths, 1,050–1,170 Å. We considered the corresponding region (that is, $\lambda/4$) in the He II forest, which for HS1700 + 64 is observed at 982–1,095 Å. Comparing the data with the modelled flux over this range we find $D_{A, \text{He II}} = 0.634 (\pm 0.027)$. This gives an effective opacity $\tau_{\text{He II}} = -\ln(1 - D_{A, \text{He II}}) = 1.00 (\pm 0.07)$ at a mean absorption redshift $z = 2.4$. The H I absorption we measure in the low-resolution MMT spectrum of HS1700 + 64

gives $D_{A, \text{H I}} = 0.20$, or $\tau_{\text{H I}} = 0.22$. (The average opacity expected¹⁵ at this redshift is $\tau_{\text{H I}} = 0.0037(1+z)^{3.46} = 0.25$.) Thus we find that at $z = 2.4$, the ratio of the He II and H I opacities produced by intergalactic gas is $\tau_{\text{He II}}/\tau_{\text{H I}} = 4.5 (\pm 0.5)$.

This ratio is determined by the mean value of the ratio of densities $\eta = \langle n(\text{He II})/n(\text{H I}) \rangle$ in the intergalactic gas, and by the distribution of the gas in space. For uniformly distributed gas, which produces the classical Gunn–Peterson effect³,

$$\tau_{\text{GP, He II}}/\tau_{\text{GP, H I}} = \eta/4$$

When the gas is clumped into regions of varying density (as is the case for the gas that produces the Lyman- α forest absorption lines), the opacity is smaller (for $\beta < 2$; see below) owing to saturation of He II lines for high column density. The reduced opacity ratio is given approximately by³

$$\tau_{\text{LF, He II}}/\tau_{\text{LF, H I}} = (\eta/4)^{\beta-1} \xi^{2-\beta}$$

where β is the index of the power-law distribution of the number of Lyman- α lines with column density ($dN_{\text{lines}}/dN_{\text{H I}} \propto N_{\text{H I}}^{-\beta}$) and ξ is the ratio of the Doppler b parameters for helium and hydrogen ($0.5 < \xi < 1.0$), which depends on the relative importance of thermal and bulk hydrodynamic velocities of the gas. Detailed calculations of absorption by simulated Lyman forest spectra (W.Z., A.F.D. and G.A.K., manuscript in preparation) show that this formula is accurate only if the column density distribution follows a power-law down to column densities where the He II lines are optically very thin, that is $N_{\text{He II}} \ll 1.6 \times 10^{14} \text{ cm}^{-2}$ for $\langle b \rangle = 30 \text{ km s}^{-1}$. Otherwise it overestimates the opacity ratio significantly, again due to the effects of He II line saturation for higher column densities. Assuming $\xi = 1.0$ (which applies if hydrodynamic motions dominate the line widths, and also maximizes the He II absorption produced by the forest lines) and taking¹⁶ $\beta = 1.5$, our measured opacity ratio requires $\eta = 80 (+20, -16)$. The Lyman- α forest would have to extend down to $N_{\text{H I}} \ll 2 \times 10^{12} \text{ cm}^{-2}$ to make the He II lines optically thin and produce the observed opacity ratio. (Truncating the distribution at $2 \times 10^{12} \text{ cm}^{-2}$ would produce an opacity ratio of only 3.) Recent observations^{16,17} of several other quasars at somewhat higher z with the Keck telescope show a Lyman forest distribution with $\beta = 1.5$ extending down to $N_{\text{H I}} = 2 \times 10^{12} \text{ cm}^{-2}$, where line detection becomes confusion-limited. For such a distribution the largest contribution to the He II opacity would come from the lines with $N_{\text{H I}}$ close to this lower limit, as higher column density lines are saturated.

An important consequence of the ionization state of the intergalactic gas is that He II absorption is sensitive primarily to conditions in the lowest-density regions of the IGM¹⁸. If the Lyman forest line distribution cannot be extrapolated to $N_{\text{H I}} \ll 2 \times 10^{12} \text{ cm}^{-2}$ with $\beta = 1.5$, then the addition of some diffuse gas producing a classical Gunn–Peterson effect could

explain a significant part of the He II opacity we measure. Detailed calculations (W.Z., A.F.D. and G.A.K., manuscript in preparation) show that the He II opacity associated with the observed Lyman- α forest lines¹⁹ in HS1700 + 64, which have $N_{\text{H I}} > 3 \times 10^{13} \text{ cm}^{-2}$, is < 0.4 for $\eta < 1,000$. Therefore, most of the He II absorption we measure arises in intergalactic gas that was, until recently, not detectable via its H I absorption.

Under the usual assumption that intergalactic gas is photoionized and optically thin at the H I and He II ionization thresholds, η is determined by the spectrum of ionizing radiation. Many theoretical investigations of the ultraviolet background spectrum^{20–22}, assuming quasars produce the radiat, and taking into account the filtering of the spectrum by the Lyman- α clouds, predict values in the range $30 < \eta < 100$. If there is also a significant contribution from young galaxies with a soft ionizing spectrum, this leads to much larger²⁰ values of η . The value $\eta = 80$ found here from a direct comparison of the He II and H I opacities provides strong support for the hypothesis that intergalactic gas at $z = 2.4$ is photoionized by radiation whose spectrum is similar to that expected from quasars. We note that if $\eta = 45$, as recently predicted by Haardt and Madau²², our measured opacity ratio would require $\beta = 1.6$ for $\xi = 1$ in the case where the absorption is attributed to an extrapolation of the Lyman- α forest to very low column densities.

The mean intensity of the ionizing background required to explain the H I and He II opacities depends on the density of the intergalactic gas. Recent calculations^{23–25} of the properties of the IGM in a universe dominated by cold dark matter and undergoing gravitational collapse from primordial fluctuations have been remarkably successful in producing the observed characteristics of the Lyman- α forest. In this theory, the weakest observable Lyman forest absorption features represent small fluctuations in the density of the IGM. These calculations predict that most of the baryons are in the intergalactic gas (including the Lyman- α forest and the lower density regions or voids) at $z = 2.4$; they also suggest^{24,25} that the mean baryon density of the Universe is at the high end of the range of values allowed by standard Big Bang nucleosynthesis theory and the observed abundances of the light elements (or perhaps even a little higher), while the ultraviolet background intensity must be at the low end of estimates from the quasar luminosity function and from the proximity effect. Our measurement of the He II opacity at $z = 2.4$ is consistent with the predictions²⁴ of this new work and will help to constrain the baryon density and the ionizing background intensity and spectrum in future simulations. The higher opacity, $\tau_{\text{He II}} > 1.7$, found⁴ at $z \approx 3.2$ is not incompatible with our result in view of the rapid increase of $\tau_{\text{H I}}$ with redshift^{15,31} and the likelihood of a similar behaviour for $\tau_{\text{He II}}$. However, this larger value may not be typical for $z \approx 3.2$, as it was obtained at a redshift close to that of Q0302 – 003, and there is also a large void³² in the Lyman- α forest within the redshift range covered by the observation. More detailed analysis of the He II absorption in quasar spectra, including fluctuations in $\tau_{\text{He II}}$ and its evolution with z , will be extremely important in helping to illuminate the development of structure in the Universe.

Note added in proof: A recent observation of Q1935 – 69 (D. Tytler and P. Jakobsen, personal communication) has failed to confirm the earlier FOS result⁴ that $\tau_{\text{He II}} \approx 1$ near this object. The new result is $\tau_{\text{He II}} > 1.5$ at $z \approx 3.1$. ■

Received 27 June 1995; accepted 19 January 1996.

- Gunn, J. E. & Peterson, B. A. *Astrophys. J.* **142**, 1633–1636 (1965).
- Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl. Ser.* **42**, 42–48 (1980).
- Miralda-Escudé, J. *Mon. Not. R. astr. Soc.* **262**, 273–276 (1993).
- Jakobsen, P. et al. *Nature* **370**, 35–39 (1994).
- Tytler, D. et al. in *QSO Absorption Lines* (eds Bergeron, G., Meylan, G. & Wampler, J.) 289–298 (Springer, Heidelberg, 1995).
- Jakobsen, P. et al. *Astrophys. J.* **417**, 528–540 (1993).
- Davidson, A. F. et al. *Astrophys. J.* **392**, 264–271 (1992).
- Davidson, A. F. *Science* **257**, 327–334 (1993).
- Reimers, D. et al. *Astr. Astrophys.* **218**, 71–77 (1989).
- Vogel, S. & Reimers, D. *Astr. Astrophys.* **294**, 377–410 (1995).
- Kruk, J. W. et al. *Astrophys. J.* **454**, L1–L6 (1995).
- Zheng, W. & Davidsen, A. F. *Astrophys. J.* **440**, L53–L56 (1995).
- Giroux, M. L., Fardal, M. A. & Shull, M. J. *Astrophys. J.* **451**, 477–483 (1995).
- Oke, J. & Korycanski, D. G. *Astrophys. J.* **255**, 11–19 (1982).
- Press, W. H., Rybicki, G. B. & Schneider, D. P. *Astrophys. J.* **414**, 64–81 (1993).
- Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A. & Rauch, M. *Astr. J.* **110**, 1526–1543 (1995).
- Songaila, A., Hu, E. M. & Cowie, L. L. *Nature* **375**, 124–126 (1995).
- Reisenegger, A. & Miralda-Escudé, J. *Astrophys. J.* **449**, 476–487 (1995).
- Rodríguez-Pascual, P. M. et al. *Astrophys. J.* **448**, 575–588 (1995).
- Miralda-Escudé, J. & Ostriker, J. P. *Astrophys. J.* **350**, 1–22 (1990).
- Madau, P. & Meiksin, A. *Astrophys. J.* **433**, L53–L56 (1994).
- Haardt, F. & Madau, P. *Astrophys. J.* (in the press).
- Cen, R., Miralda-Escudé, J., Ostriker, J. P. & Rauch, M. *Astrophys. J.* **437**, L9–L12 (1994).
- Miralda-Escudé, J., Cen, R., Ostriker, J. P. & Rauch, M. *Astrophys. J.* (in the press).
- Hemquist, L., Katz, N., Weinberg, D. H. & Miralda-Escudé, J. *Astrophys. J.* **457**, L51–L56 (1996).
- Cardelli, J., Clayton, G. & Mathis, J. *Astrophys. J.* **345**, 245–256 (1989).
- Burstein, D. & Heiles, C. *Astrophys. J.* **225**, 40–55 (1978).
- Stark, A. A. et al. *Astrophys. J. Suppl. Ser.* **79**, 77–104 (1992).
- Eastman, R. G. & MacAlpine, G. M. *Astrophys. J.* **299**, 785–798 (1985).
- Netzer, H., Elitzur, M. & Ferland, G. *Astrophys. J.* **299**, 752–762 (1985).
- Jenkins, E. B. & Ostriker, J. P. *Astrophys. J.* **376**, 33–42 (1991).
- Dobrzycki, A. & Bechtold, J. *Astrophys. J.* **377**, L69–L72 (1991).

ACKNOWLEDGEMENTS. We thank our colleagues on the Hopkins Ultraviolet Telescope Team whose efforts have made this work possible. We also thank the many individuals at NASA and elsewhere who helped made the Astro-2 mission a success. This work was supported by NASA.

X-ray holography with atomic resolution

Miklós Tegze & Gyula Faigel

Research Institute for Solid State Physics, H-1525 Budapest, POB 49, Hungary

DIFFRACTION methods for crystallographic structure determination suffer from the so-called ‘phase problem’; a diffraction pattern provides intensity but not phase information for the scattered beams, and therefore cannot be uniquely inverted to obtain the crystal structure of a sample. Holographic methods¹, on the other hand, offer a means of extracting both intensity and phase information. To be useful for crystallographic applications, holography must be implemented with radiation of sufficiently small wavelength to resolve atomic-scale features². One method, electron-emission holography^{3–9}, uses electron waves and is a powerful tool for studying surface structure; but it cannot image the internal structure of solids because of complications arising from the highly anisotropic nature of electron scattering processes. A proposed alternative method uses X-rays^{2,10–13}, which scatter more isotropically than electrons. Here we demonstrate the efficacy of atomic-scale X-ray holography by obtaining direct images of the three-dimensional arrangement of strontium atoms in the cubic perovskite SrTiO₃. With more intense synchrotron sources for illumination, and with the development of improved X-ray detectors, X-ray holography should become a powerful general technique for unambiguous structure determination in condensed matter systems.

Holography with ‘hard’ X-ray photons is based on the ‘inside’ source concept². In practice, these inside sources are the atoms or nuclei in the sample. How can we force atoms to emit short-wavelength X-ray photons? Using an external X-ray source, core-level electrons can be removed leaving behind a hole. In the relaxation process, a photon is emitted—called fluorescent X-radiation—with well defined energy. One can detect these photons far from the sample. The photon can reach the detecting surface directly (without interacting with the atoms in the sample), or it can be elastically scattered by the neighbouring atomic electrons. (The inelastic scattering is small compared to elastic scattering, and gives a smooth background). Interference between these two processes results in an angle-dependent intensity, which is the hologram itself. Nuclei can also serve as sources of γ -radiation. There are nuclei having low-lying excited states (10–

100 keV). (These isotopes are often called Mössbauer isotopes because many of them are used for Mössbauer studies¹⁴.) These nuclear states can be populated in two ways: exciting the isotopes by electromagnetic radiation, or by having radioactive parent nuclei in the sample which decay via the excited state in question. The formation of a γ -ray hologram is analogous to the atomic electron excitation case except that the resonant nuclear scattering has to be taken into account as well as the electronic scattering. In our demonstration experiment fluorescent photons were used, so we restrict the discussion to this case. The mathematical foundation of the imaging process and reconstruction has been published elsewhere^{10,15,16}. Here we give the basic formula only, to help understanding of the experimental difficulties which will be discussed in more detail. The measured intensity I is given by

$$I = I_0/R^2 \left[1 + \left| \sum a_j \right|^2 + 2\text{Re} \left(\sum a_j \right) \right] \quad (1)$$

where I_0 is the intensity of the source atom, a_j is the wave amplitude scattered by the j th atom, R is the sample-detector distance and Re stands for the real part. The first term in the square bracket corresponds to the intensity of the direct or reference beam (R) leaving the sample without scattering. The second term is the intensity of the scattered waves—in holography they are called object waves (O)—while the third component is an interference term between the reference and object waves (RO). (Here we include only single scattering events. In the case of multiple scattering the form of equation (1) does not change¹⁵. Further, for X-rays the scattering cross-section is small compared to that of electrons, and the multiple scattering contribution can be neglected in most cases.) This last sum is the hologram. In a measurement all three components are simultaneously detected, and the holographic information has to be separated in the data analysis. If the scattering of the emitted photons is weak, the amplitude of the O wave is small ($\sim 10^{-6}$) compared to that of the RO wave ($\sim 10^{-3}$), so it can be neglected. (The values of the different terms were estimated for an element with medium atomic number.) The direct beam is angle-independent, so in principle it can be easily subtracted from the measured intensity.

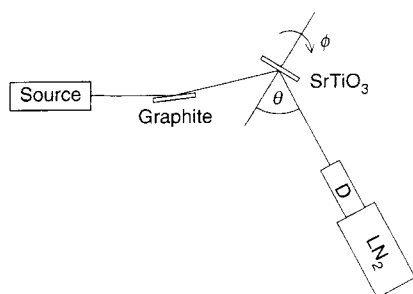
Until now, we have considered the case where only one atom is emitting photons. But in practice a macroscopic sample is used and this makes the separation of the first and second terms much more difficult. In a macroscopic sample many atoms radiate independently and separate holograms are formed. In order to get a single hologram, all radiating atoms have to have the same environment oriented in the same way and the size of the sample has to be small. (The size of the crystal or the illuminated area has to be smaller than $R\lambda/r$ where R , λ and r correspond respectively to the detector-sample distance, the wavelength of the fluorescent radiation and the typical atomic distances to be resolved.) These conditions are satisfied by a small single-crystal specimen. However, in this case the second term in equation (1) might take substantial values in well defined directions owing to the translational symmetry of the sample. In our earlier work it was shown

that the spatial frequency of this component is much larger than that of the holographic information given by the nearest-neighbour environment of the radiating atom and therefore it can be removed by a low-pass filter¹⁰. The second problem which results from the macroscopic nature of the sample comes from a combined effect of the shape and the absorption. A sample of negligible absorption would be ideal. In reality this is not the case. Photons originating from different places in the sample have different probabilities of absorption before leaving the specimen. Therefore a correction of the measured intensity is necessary, which depends on the shape of the sample and the geometry of the experiment. We used a flat specimen and the geometry shown in Fig. 1. In this arrangement, the intensity decreases to zero monotonically by varying θ from 0° to 90° . In a practical range of $0^\circ < \theta < 60^\circ$, the intensity may change by 50%. It is clear that the proper correction of this large variation is crucial to see the expected 0.3% oscillation in the intensity of the hologram. The third major factor influencing the measured intensity is detector-related. The detectors have a dead time, that is, a time interval following an event in which the detector cannot count a new photon. As the count rate increases, more and more events are not detected. This dead time has to be corrected precisely. For a single channel (one point of the hologram) all the above corrections could be done in a straightforward way, but for the full two-dimensional surface of the hologram ($\sim 2,400$ pixel in this case) a correct subtraction is difficult. We list briefly what we have done to avoid the distorting effects of the above factors.

(1) The theoretical form of the absorption correction was calculated for a flat specimen. (2) An energy integrated count rate was measured together with the fluorescent photons to correct for the dead time. (3) a low-pass filter was applied to reduce the O wave. (4) A high-pass filter was applied to reduce noise. (5) A carefully planned experimental arrangement and multiple scanning were used to ensure long-term stability and reproducibility.

The demonstration experiment was done on a SrTiO_3 single-crystal platelet 30 mm in diameter and 0.5 mm thick. In the perovskite-type SrTiO_3 structure (lattice parameter $a = 3.9 \text{ \AA}$) the Sr atoms are arranged on a simple cubic lattice. The surface of the crystal was almost parallel to the (100) plane. Because the Sr K absorption edge is at 16.105 keV, the K electrons can be excited by Mo K_α radiation ($\sim 17,425 \text{ eV}$). The 2,402-pixel hologram, after correcting for the absorption and the direct beam is depicted in the top panel of Fig. 2. At least 4×10^6 counts were collected in every pixel, and the overall anisotropy in these measured intensities was 0.3%, in agreement with theoretical expectations. The imperfect four-fold symmetry of the hologram is caused by a tilt of the physical surface of the crystal from the (100) atomic plane by $\sim 3^\circ$; this tilt also causes the strong central peak in the hologram to be off-centre to the left. (See Fig. 2 top panel.) The reconstruction was done using the Helmholtz-Kirchhoff formula¹⁵ and applying the corrections described above. Three atomic planes parallel to the crystal physical surface are shown in three dimensions in the bottom panel of Fig. 2. Only the Sr atoms can be seen in the figure because the atomic scattering factor of Sr is about twice that of Ti, and the scattering of oxygen is even smaller. Note that the atoms are not dots but that they have different sizes and shapes in the different planes. Also, some atoms are not centred exactly at the expected positions. These effects are caused by the geometry of our experiment and by the appearance of the twin image. The twin image is known from traditional holography. It appears as a mirror image of the original object, usually spatially separated from the real image. However, in our case each atom has its mirror counterpart owing to the centrosymmetry of the lattice. Therefore, the twin and real images of different atoms are at the same place. As the phases of these two waves vary differently, the interference between them can result in an intensity modulation close to atomic positions which can shift the apparent centres of some atoms¹⁰, or even lead to a cancellation of a given atomic image¹². In electron holography efforts have been made to solve this

FIG. 1 Schematic view of the experimental apparatus. The graphite monochromator selected Mo K_α radiation that was incident on the sample at $\sim 45^\circ$. The two-dimensional hologram measured by turning the sample about ϕ and moving the Ge solid-state detector in the plane determined by the incident beam and the ϕ axis, so as to vary angle θ . A standard 2 kW X-ray generator with a sealed tube was used.



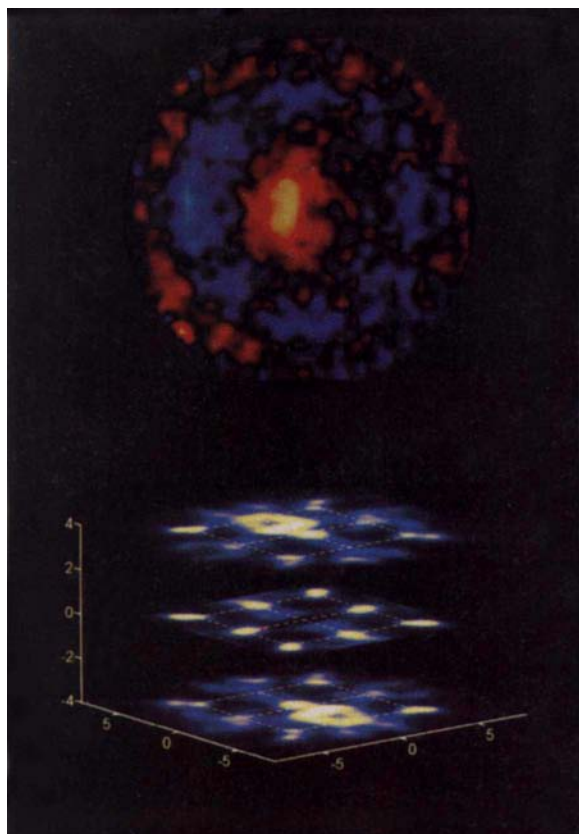


FIG. 2 Top panel, the hologram of SrTiO₃ after absorption correction and subtracting the contribution of the direct beam. This image is a projection of the spherical coordinates on the k_x - k_y (horizontal, vertical) plane ($k_x = \sin \theta \sin \phi$, $k_y = \sin \theta \cos \phi$; the definitions of θ and ϕ are given in Fig. 1). The radius of the hologram is 0.92. Bottom panel, three-dimensional holographic image of SrTiO₃, showing only the Sr atoms (see text for details). The outline of eight elementary cells are shown by solid lines. The scale is given in ångströms.

problem, and very promising methods have been developed^{12,17-19}. Here we do not use them as we intend to demonstrate that, without any *a priori* knowledge and a complicated evaluation procedure, the three-dimensional atomic order can be deduced. The different size of the atoms may be explained by the flat plate geometry. This geometry leads to different resolution in the plane which contains the source atom and is parallel with the crystal physical surface (in-plane), and in planes perpendicular to the crystal surface (out-of-plane). Because the diameter of the in-plane projection of the hologram is close to the maximum value of $2\pi/\lambda$, the in-plane resolution is ~ 0.7 Å as expected from the wavelength of the fluorescent radiation. The out-of-plane resolution (almost 2 Å) is determined by the size of the projection of the hologram to the direction perpendicular to the crystal surface, which is only about $(1/3)(2\pi/\lambda)$. Besides degrading the resolution, this also produces spurious oscillations in the reconstructed image. This problem could be overcome by using a small spherical sample instead of a flat one, and by recording the hologram on the full sphere. In this case the resolution would be the same in all directions. The effect of the above factors occur together and determine the final shape and size of the atoms in the reconstructed image.

We now consider future possibilities. Use of much brighter synchrotron sources for excitation, together with the development of fast one-dimensional high energy resolution detectors will allow the collection of much higher-quality data sets in a much shorter time. Therefore, in many cases our method might become a practical tool in solving the phase problem of single-crystal

diffraction. A further step would be the development of advanced X-ray cameras with good energy resolution. Such cameras could follow, in real time, the dynamics of processes such as the change of atomic positions in a solid-solid phase transition. Our technique may also involve nuclear “inside” sources and scatterers. In a solid, excited nuclei can emit photons without recoil¹⁴ and with very well defined energy. These photons are elastically scattered not only by the electrons of the neighbouring atoms but also by the nuclei. The resonant nuclear scattering depends on the hyperfine fields in the sample. Therefore, one could obtain a three-dimensional picture of magnetic order or distinguish crystallographically equivalent atoms which sense non-equivalent electric field gradients or charge states. Further, the cross-section of nuclear scattering can be higher than that of the electronic charge scattering, resulting in an increased, isotope-selective sensitivity. □

Received 31 July 1995; accepted 24 January 1996.

- Gabor, D. *Nature* **161**, 777–778 (1948).
- Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Applications* (eds Attwood, D. T. & Boker, J.) 361–467 (AIP Conf. Proc. No. 147, American Institute of Physics, New York, 1986).
- Harp, G. R., Saldin, D. K. & Tonner, B. P. *Phys. Rev. Lett.* **65**, 1012–1015 (1990).
- Thevuthasan, S. et al. *Phys. Rev. Lett.* **70**, 595–598 (1993).
- Terminello, L. J., Barton, J. J. & Lapiano-Smith, D. A. *Phys. Rev. Lett.* **70**, 599–602 (1993).
- Li, H., Tong, S. Y., Naumovic, D., Stuck, A. & Osterwalder, J. *Phys. Rev.* **B47**, 10036–10039 (1993).
- Saldin, D. K., Harp, G. R. & Chen, X. *Phys. Rev.* **B48**, 8234–8244 (1993).
- Han, Z. L. et al. *Surf. Sci.* **258**, 313–327 (1991).
- Wei, C. M., Tong, S. Y., Wedler, H., Mendez, M. A. & Heinz, K. *Phys. Rev. Lett.* **72**, 2434–2437 (1994).
- Tegze, M. & Faigel, G. *Europhys. Lett.* **16**, 41–46 (1991).
- Maalouf, G. J., Hoch, J. C., Stern, A. S., Szöke, H. & Szöke, A. *Acta crystallogr.* **A49**, 866–871 (1993).
- Len, P. M., Thevuthasan, S., Fadley, C. S., Kaduwela, A. P. & Van Hove, M. A. *Phys. Rev.* **B50**, 11275–11278 (1994).
- Szöke, A. *Acta crystallogr.* **A49**, 853–866 (1993).
- Mössbauer, R. L. Z. *Phys.* **151**, 124–143 (1958); *Naturwissenschaften* **45**, 538–539 (1958).
- Barton, J. J. *Phys. Rev. Lett.* **61**, 1365–1359 (1988).
- Saldin, D. K. & de Andres, P. L. *Phys. Rev. Lett.* **64**, 1270–1273 (1990).
- Chen, X. & Saldin, D. K. *Phys. Rev.* **B50**, 17463–17470 (1994).
- Barton, J. J. *Phys. Rev. Lett.* **67**, 3106–3109 (1991).
- Tong, S. Y. & Huang, H. & Wei, C. M. *Phys. Rev.* **B46**, 2452–2459 (1992).

ACKNOWLEDGEMENTS. This work was supported by OTKA and PHARE ACCORD.

Rapid climate changes in the tropical Atlantic region during the last deglaciation

Konrad A. Hughen*, Jonathan T. Overpeck*†, Larry C. Peterson‡ & Susan Trumbore§

* INSTAAR and Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309, USA

† NOAA Paleoclimatology Program, NGDC, Boulder, Colorado 80303, USA

‡ RSMAS, University of Miami, Miami, Florida 33149, USA

§ Department of Earth Systems Science, University of California, Irvine, California 92717, USA

THE climate system is capable of changing abruptly from one stable mode to another¹⁻³. Rapid climate oscillations—in particular the Younger Dryas cold period during the last deglaciation—have long been recognized from records throughout the North Atlantic region⁴⁻¹⁴, and the distribution of these records at mostly high latitudes suggests that the changes were caused by rapid reorganizations of the North Atlantic thermohaline circulation^{6,8,10,15}. But events far from the North Atlantic region that are synchronous with the Younger Dryas¹⁶⁻¹⁹ raise the possibility that a more global forcing mechanism was responsible²⁰. Here we present high-resolution records of laminated sediments of the

last deglaciation from the Cariaco basin (tropical Atlantic Ocean) which show many abrupt sub-decade to century-scale oscillations in surface-ocean biological productivity that are synchronous with climate changes at high latitudes. We attribute these productivity variations to changes in or duration of upwelling rate (and hence nutrient supply) caused by changes in trade-wind strength, which is in turn influenced by the thermohaline circulation through its effect on sea surface temperature^{6,21}. Abrupt climate changes in the tropical Atlantic during the last deglaciation are thus consistent with a North Atlantic circulation forcing mechanism.

The Cariaco basin (10° 40' N, 65° W) is an anoxic marine basin located within the trade-wind belt off the northern coast of Venezuela (Fig. 1). The climate of the Cariaco basin region has a large seasonal signal, controlled by the annual north-south migration of the Intertropical Convergence Zone (ITCZ) in the Atlantic Ocean. When the ITCZ is farthest south (winter), rainfall in the Cariaco basin area is at a minimum and strong trade winds blow along the coast, inducing Ekman upwelling. When the ITCZ moves north (summer), the trade winds and upwelling diminish and the rainy season occurs. This results in a seasonal alternation of a dry, upwelling season and a rainy, non-upwelling season²¹⁻²³. Sedimentation patterns within the Cariaco basin accurately follow the climate, with annual deposition of a lighter-coloured, plankton-rich layer, followed by a darker-coloured, terrigenous grain-rich layer. Sediment cores recovered from the basin have continuously laminated sediments from the surface to a depth radiocarbon-dated at 12.6 kyr before present (BP). Previous work with ²¹⁰Pb and historical records has confirmed that the laminae couplets over the last century are annually deposited varves²³.

The construction of a multiple-core varve chronology and laminae thickness records is time-consuming, and is still in progress. In order to construct more rapidly a continuous record of Cariaco basin variability, we measured relative reflectivity (grey scale) on fresh surfaces of four cores from near the centre of the basin (Fig. 2). Radiocarbon age control through the laminated interval is provided by 20 accelerator mass spectrometry (AMS) ¹⁴C dates on monospecific samples of the planktonic foraminifer *Globigerina bulloides* from one of the cores (Fig. 2). A reservoir correction age of 420 years was calculated for the Cariaco basin using two ¹⁴C dates on near-surface sediments of precisely known age²³. This reservoir correction may have changed through time as a result of changes in upwelling and trade-wind strength. However, these two factors influence reservoir age in opposite directions, and the present value is close to that of the non-upwelling, open Atlantic (~400 years), indicating the changes may have been small. Detailed correlations between the cores, independent of the grey-scale data, were made using individual 'marker' laminae observed in photographs, X-radiographs and petrographic thin sections. The cores display nearly identical grey-scale records, even in fine detail, indicating that the changes reflect variable surface conditions over the Cariaco basin in general, and are not limited to single locations. Comparing grey scale to a preliminary thickness record for the light-coloured, upwelling-season laminae from two of the cores shows that grey-scale values correlate well with average light-lamina thickness (Fig. 3). As light laminae increase in thickness, the sediments are more reflective and appear lighter, and vice versa. Slight disagreements in detail between the records may be due to the preliminary state of the lamina thickness record, which consists of 1-cm averages over much of its length. In general the agreement is good, suggesting that grey scale is controlled by the annual amount of biogenic productivity.

High-deposition-rate palaeoclimate records show that abrupt deglacial changes occurred across broad

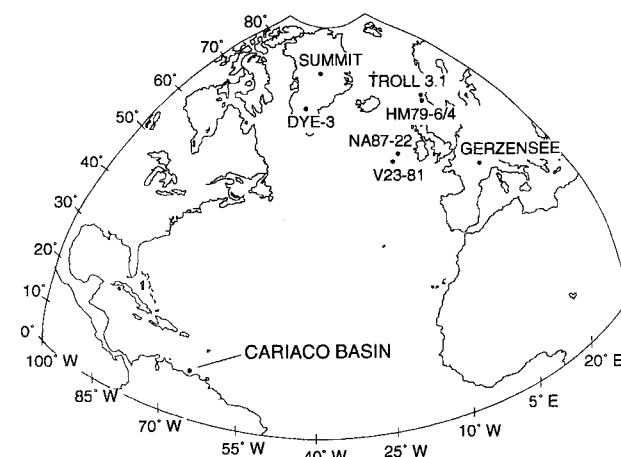


FIG. 1 Map of the North Atlantic region showing the location of the Cariaco basin in the western tropical Atlantic and the locations of high-latitude sites (referred to in the text) with records that correlate well with the productivity record of the Cariaco basin.

regions of the high-latitude North Atlantic. Dansgaard *et al.*⁹ showed synchronous changes of oxygen isotope ($\delta^{18}\text{O}$) records from the Dye-3 ice core in southern Greenland and calcite from Lake Gerzensee, Switzerland. Lehman and Keigwin¹⁰ correlated an abundance record of the polar foraminifer *Neogloboquadrina pachyderma* (sinistral) from marine core Troll 3.1 off the Norwegian coast to another *N. pachyderma* (s.) record from core V23-81 (Fig. 1), and correlated these with the Dye-3 and Gerzensee records. Koç-Karpuz and Jansen¹² showed similar deglacial changes in a spliced sea surface temperature (SST) record based on diatom assemblages from two marine cores, HM79-4 and -6, near core Troll 3.1 (Fig. 1). The same pattern of oscillations is repeated in $\delta^{18}\text{O}$ from ice cores at other Greenland sites, including Camp Century, Renland and Summit¹⁴⁻²⁶, as well as in snow accumulation, concentrations of several chemical species, and electrical conductivity from the GISP2 Greenland ice-core^{11,27,28}. Changes of the same character are also evident in $\delta^{18}\text{O}$ records

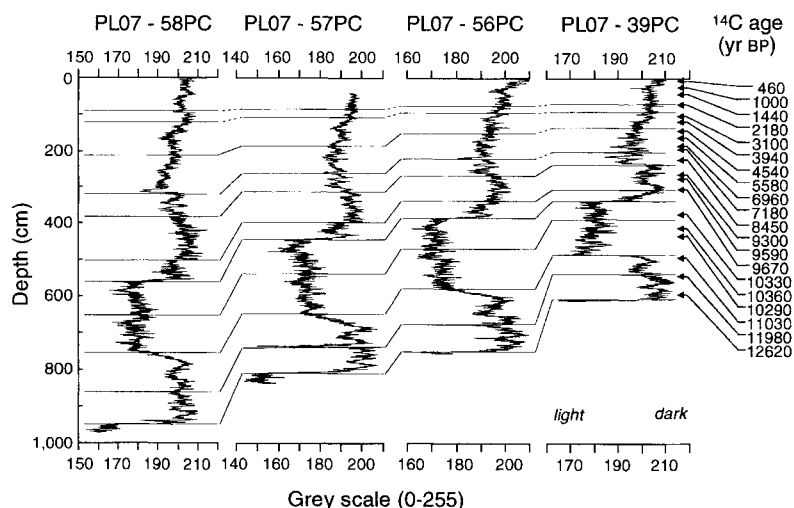


FIG. 2 Relative reflectivity (grey scale) from Cariaco basin cores PL07-58PC, -57PC, -56PC and -39PC. The correlation lines are based on individual laminae recognizable from photographs, X-radiographs and petrographic thin sections, and are intended as a check on grey-scale correlations. Age control is provided by 20 AMS ¹⁴C dates on monospecific samples of the planktonic foraminifer *Globigerina bulloides* from core PL07-39PC, using a 420-year marine reservoir correction. The detailed pattern of grey-scale variability is clearly repeated in all cores. These data are archived at the World Data Center-A for Paleoclimatology (NOAA, Boulder, Colorado).

from the Faulensee (near Gerzensee) and Chirens bogs in the Swiss and French alps⁵, and in pollen curves from Lake Sandvikvatn, Norway⁷.

Together, these records clearly demonstrate that large areas of the high-latitude North Atlantic region experienced similar abrupt changes during the last deglaciation. The decade- to century-scale oscillations identified in these records correlate well with the grey-scale record of the tropical Cariaco basin (Fig. 4a). The Bølling/Allerød interval (14,500–12,500 calendar years BP, 14.5–12.5 cal. kyr BP)²⁹, for example, was interrupted by three decade- to century-scale events in each of these records. The events, from oldest to youngest, are named here the inter-Bølling cold period (IBCP), the Older Dryas (OD), and the inter-Allerød cold period (IACP), after refs 10 and 12. The Younger Dryas (12.5–11 cal. kyr BP) was an ~1,500-year-long interval of higher productivity in the Cariaco basin marked by abrupt transitions. Pollen records from the same latitude in Costa Rica indicate that temperature dropped by 2–3 °C at that time³⁰. The termination of the Younger Dryas in the Cariaco basin, estimated from thin sections of the varved sediments, occurred in less than a decade, in agreement with snow accumulation and other annual records from Greenland^{11,27,28}. Following the Younger Dryas termination was an oscillation during the Preboreal chron (PB) at 11 cal. kyr BP. After about 1,000 years, core PL07-56PC shows another series of oscillations, called early Holocene (EH), I, II and III, from 9 to 7.5 cal. kyr BP (Fig. 4a). EH III is clearly discernable in the Greenland record and changes at this time are seen in reconstructed sea surface conditions from foraminifera in North Atlantic core NA87-22 (ref. 14) and other records throughout the Atlantic basin^{21,31–34}. The Cariaco basin record shows continued variability throughout the rest of the Holocene, after 8 cal. kyr BP (Figs 2 and 4a).

The strong similarities between the Cariaco basin and GRIP records at the decade- to century-scale prompted us to compare

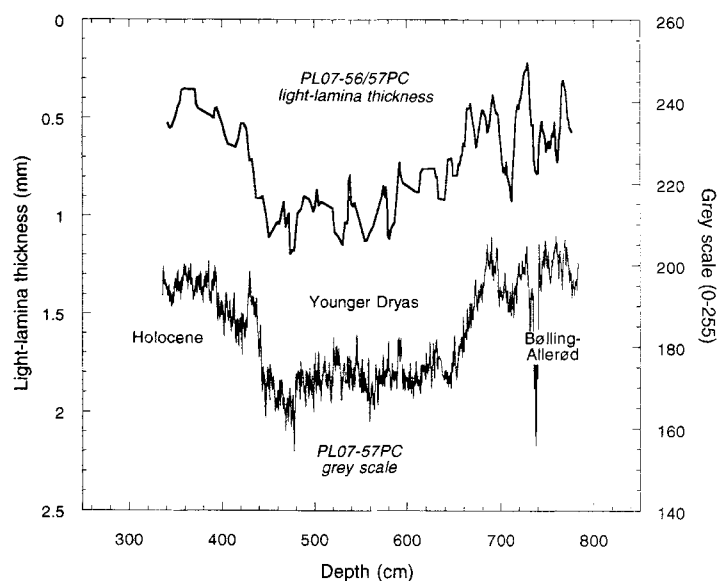


FIG. 3 Comparison of a preliminary spliced light-lamina-thickness record from cores PL07-56PC and -57PC to grey scale from core PL07-57PC. The thickness measurements, everywhere except during the Bølling/Allerød where they are annual (see Fig. 4b), are averages over 1-cm intervals and were made ~2–15 cm apart, depending on where thin sections were available for analysis. The large-scale excursions to brighter (low) values seen in the grey-scale record clearly coincide with increased average thickness of light laminae, and hence provide a proxy for changes in biogenic productivity.

these records at higher resolution. We compared a completed light-lamina thickness record from PL07-57PC to $\delta^{18}\text{O}$ data from GRIP over an interval including the inter-Allerød cold period and Older Dryas (Fig. 4b). Both records are constrained by annual chronologies, although the light-lamina record is smoothed with a 15-year running average and the GRIP core samples at this depth average 13–14 years together. The most negative $\delta^{18}\text{O}$ excursions seen in the GRIP record lasted approximately 131 and 27 years for the IACP and OD, respectively. The comparable events in the Cariaco basin were of similar duration, 127 and 21 years. In addition to the chronological agreement, there is also considerable similarity in the decade-scale patterns of variability in both records. Given the geographical distance separating central Greenland from the southern Caribbean Sea, the close match of the pattern and duration of decadal events between the two records is striking.

The similarity of detail between Cariaco basin and high-latitude North Atlantic records suggests a common forcing mechanism.

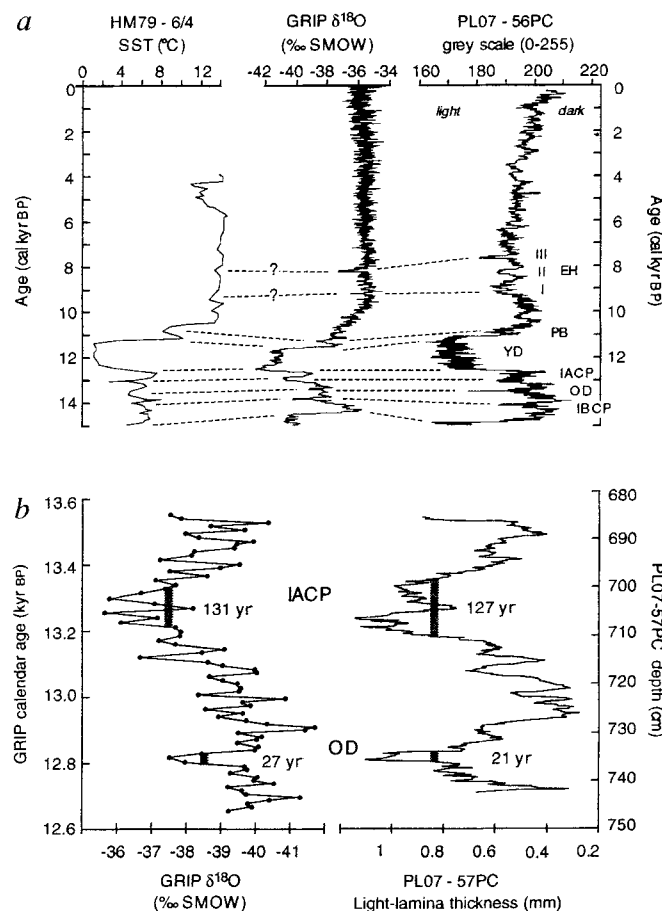


FIG. 4 a, Comparison of grey scale from Cariaco basin core PL07-56PC to $\delta^{18}\text{O}$ in the GRIP ice core²⁴ and August SST reconstructed from diatoms in lower-resolution marine cores HM79-6 and -4 (ref. 12). The Cariaco basin and HM79-6/4 records were converted to calendar ages for comparison to GRIP using the CALIB calibration program of Stuiver and Reimer²⁹. The comparison shows very similar patterns and timing of changes from 14–8 cal. kyr BP. Names of events are described in the text. b, Detailed comparison of $\delta^{18}\text{O}$ from the GRIP ice core²⁴ with changes in a continuous sequence of light-lamina thickness measurements from core PL07-57PC. Both records are constrained by annual chronologies, although neither record is sampled at annual resolution. The interval of comparison includes the inter-Allerød cold period (12.9–13 cal. kyr BP) and Older Dryas (13.4 cal. kyr BP) events (IACP and OD from a). The durations of the two events, measured independently in both records, are very similar, as is the detailed pattern of variability at the decadal timescale.

The brevity of the Older Dryas and other short events in Fig. 4b implies that the events occurred at nearly the same time. If there were a significant time lag (more than one wavelength of the briefest event) and signals would probably have been attenuated. The atmosphere has the geographic range and rapid response to link sub-decade changes in widely separated regions. Cariaco basin grey scale and light-lamina thickness record increases in primary productivity at times of cold in the North Atlantic. The most likely way to have increased productivity is to have increased intensity or duration of upwelling, in this case through strengthening of the trade winds. General circulation model results suggest that a reduction in North Atlantic SST would result in increased annually averaged trade winds over the tropical North Atlantic^{6,21}. This would cause an increase in Cariaco basin upwelling and productivity whenever North Atlantic SST was lowered. Thus it appears that shifts in North Atlantic thermohaline circulation, in addition to causing observed changes in high-latitude records, could have influenced trade-wind strength in the tropical Atlantic and caused the productivity record in the Cariaco basin.

The North Atlantic may have played a prominent role in controlling climate over most of the tropical Atlantic during the last deglaciation. Lake level records from northwestern Africa show periods of drought at about 11–10 and 8 ¹⁴C kyr BP (refs 31–33), corresponding to the YD and EH intervals in the North Atlantic and Cariaco basin. These periods in the African records have been attributed partly to the influence of a cold North Atlantic on North African summer temperatures and monsoon precipitation^{31–33}. The similar timing of large changes in high- and low-latitude records may indicate a coupled tropical/high-latitude North Atlantic climate system operating during the last deglaciation until about 8 ¹⁴C kyr BP. The mechanism for this coupling would be changes in North Atlantic SST influencing tropical Atlantic trade wind and summer monsoon strength. □

Received 24 July 1995; accepted 12 January 1996.

- Manabe, S. & Stouffer, R. J. *J. Clim.* **1**, 841–866 (1988).
- Broecker, W. S. & Denton, G. H. *Geochim. cosmochim. Acta* **53**, 2465–2501 (1989).
- Stocker, T. F. & Wright, D. G. *Nature* **351**, 729–732 (1991).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145–214 (1981).
- Siegenthaler, U., Eicher, U., Oeschger, H. & Dansgaard, W. *Ann. Glaciol.* **5**, 149–152 (1984).
- Rind, D., Peteet, D., Broecker, W. S., McIntyre, A. & Ruddiman, W. *Clim. Dyn.* **1**, 3–33 (1986).
- Paus, A. *Boreas* **17**, 113–139 (1987).
- Broecker, W. S. et al. *Paleoceanography* **3**, 1–19 (1988).
- Dansgaard, W., White, J. W. C. & Johnsen, S. J. *Nature* **339**, 532–534 (1989).
- Lehman, S. J. & Keigwin, L. D. *Nature* **356**, 757–762 (1992).
- Alley, R. B. et al. *Nature* **362**, 527–529 (1993).
- Koc-Karpuz, N. & Jansen, E. *Paleoceanography* **7**, 499–520 (1992).
- Bond, G. et al. *Nature* **365**, 143–147 (1993).
- Duplessy, J. C. et al. *Nature* **358**, 485–488 (1992).
- Broecker, W. S. et al. *Paleoceanography* **5**, 469–477 (1990).
- Linsley, B. K. & Thunell, R. C. *Paleoceanography* **5**, 1025–1039 (1990).
- Peteet, D. et al. *Eos* **75**, 587–590 (1993).
- Denton, G. H. & Hardy, C. H. *Science* **264**, 1434–1437 (1994).
- Thompson, L. G. et al. *Science* **269**, 46–50 (1995).
- Broecker, W. S. *Nature* **372**, 421–424 (1995).
- Overpeck, J. T., Peterson, L. C., Kipp, N., Imbrie, J. & Rind, D. *Nature* **338**, 553–557 (1989).
- Peterson, L. C., Overpeck, J. T., Kipp, N. G. & Imbrie, J. *Paleoceanography* **6**, 99–119 (1991).
- Hughen, K. A., Overpeck, J. T., Peterson, L. C. & Anderson, R. F. in *Paleoclimatology and Paleoceanography from Laminated Sediments* (ed. Kemp, A. E. S.) (The Geological Society, London, in the press).
- Dansgaard, W. et al. *Nature* **364**, 218–220 (1993).
- Johnsen, S. J. et al. *Nature* **359**, 311–313 (1992).
- Groote, P. M., Stuiver, M., White, J. W. C., Johnsen, S. & Jouzel, J. *Nature* **366**, 552–554 (1993).
- Mayewski, P. A. *Science* **263**, 1747–1751 (1993).
- Taylor, K. C. et al. *Nature* **361**, 432–436 (1993).
- Stuiver, M. & Reimer, P. J. *Radiocarbon* **35**, 215–230 (1993).
- Islebe, G. A., Hooghiemstra, H. & van der Borg, K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **117**, 73–80 (1995).
- Street-Perrott, F. A. & Perrott, R. A. *Nature* **343**, 607–612 (1990).
- Lamb, H. F. et al. *Nature* **373**, 134–137 (1995).
- Gasse, F. & Van Campo, E. *Earth planet. Sci. Lett.* **126**, 435–456 (1994).
- Blunier, T., Chappellaz, J., Schwander, J., Stauffer, B. & Raynaud, D. *Nature* **374**, 46–49 (1995).

ACKNOWLEDGEMENTS. We thank C. Rooth, D. Murray and M. Kashgarian for valuable discussions and comments on the manuscript. We also thank H.-L. Lin for preparation of AMS ¹⁴C samples for dating, and R. Spackman and M. Duval for help in the task of making thin sections. This work was supported by the US NSF and NOAA, as well as by NSF funding to the Lamont-Doherty Earth Observatory Deep-Sea Sample Repository.

Intense mixing of Antarctic Bottom Water in the equatorial Atlantic Ocean

K. L. Polzin*, K. G. Speer†, J. M. Toole‡
& R. W. Schmitt‡

* University of Washington, School of Oceanography, Box 357940, Seattle, Washington 98195, USA

† Laboratoire de Physique des Océans, IFREMER/CNRS, 29280 Plouzané, France

‡ Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

THE spreading of Antarctic Bottom Water—the densest global-scale water mass—is highly constrained by ocean-floor topography. In the Atlantic Ocean, the Mid-Atlantic Ridge confines this water mass mainly to the western basins, the bottom waters in the eastern basins being renewed by flows through gaps in the ridge¹. One such gap is the Romanche fracture zone, a large offset of the ridge which straddles the Equator. It has been observed² that sills within this fracture zone block the passage of waters colder than ~0.9 °C; warmer, less dense waters passing over the sills appear to cascade downslope where they are modified by mixing. Here we present direct measurements which quantify these processes. The flow is vertically sheared and exhibits remarkably intense turbulence, comparable to that seen at the ocean surface in the presence of winds of ~10 m s⁻¹. This turbulence mixes the densest waters passing through the fracture zone with the warmer, overlying waters, so that the coldest waters exiting this region have been warmed by ~0.6 °C during transit. Topographic obstructions and turbulent mixing together thus determine the properties of the flows renewing the deepest waters of the Atlantic Ocean's eastern basins.

Antarctic Bottom Water (AABW) represents a primary branch of the thermohaline circulation, the name given to the buoyancy-driven flow in the ocean. Bottom waters, created at high latitude when surface waters are made dense by heat loss and evaporation to the atmosphere, spread along the sea floor from their formation areas. In thermodynamic steady state, closure of the thermohaline circulation is accomplished by vertical mixing and upwelling. Mixing with warm overlying waters heats and lowers the density of the fluid at depth so that newly formed bottom waters can intrude below, forcing upward motion. The intensity of the mixing and its geographical distribution, as well as the causal physical mechanisms, have been subjects of great speculation in oceanography. Two conceptual models have been suggested: (1) nearly uniform modest-intensity mixing supported by internal wave breaking^{3,4} and (2) vigorous mixing only in close proximity to the sea floor, with subsequent horizontal distribution of mixing products into the relatively quiescent interior^{5,6}. In the second case, bottom boundary layer formation⁶ and breaking of bottom-reflected⁷ or bottom-generated⁸ internal waves have been suggested as possible mixing mechanisms. Recent studies have indicated only low-intensity turbulent mixing in the deep-ocean interior^{9,10}, arguing against (1). Here we explore a third scenario: that significant mixing occurs at bathymetric constrictions for the abyssal circulation.

The data for this study were obtained in November and December 1994 from the French research vessel *Le Noroit*. Our principal instrumentation consisted of the High-Resolution Profiler (HRP), a free-falling, internally recording profiler¹¹, and a conventional wire-lowered Conductivity–Temperature–Depth (CTD) instrument. The CTD returns estimates of the ocean's temperature and salinity variations with depth. The HRP adds horizontal velocity estimates and turbulent temperature and velocity gradient information. HRP profiles (28) and multiple CTD lowerings were made to depths as great as 5,300 m in the Romanche fracture zone (RFZ), two HRP stations were occupied

in the Chain fracture zone and 25 profiles were obtained in adjacent regions. Work within the RFZ chiefly focused downstream from the main sill to the exit in the Sierra Leone basin (Fig. 1). High-resolution bathymetric maps¹² show that the bottom-water flow is well constrained by the northern wall of the RFZ which extends eastwards at $0.5\text{--}1.0^\circ\text{N}$ to about $13^\circ 10'\text{W}$. The southern border has numerous gaps which open into a region of broken topography with channels that reach south to the Chain fracture zone.

The distribution of potential temperature, Figure 2a, (and potential density) along the RFZ shows that the most dramatic response of the fluid is within the less-dense components of the AABW (the water mass defined as potential temperature $<1.9^\circ\text{C}$). Negligible uplift of the deepest isotherms is observed upstream of the main sill². This suggests that the densest AABW is blocked: specifically, fluid parcels colder than 0.9°C do not have sufficient initial kinetic energy to lift up and over the main RFZ sill, nor do the less-dense components of AABW accelerate sufficiently¹³ to result in an uplifting and transport of AABW colder than 0.9°C . In this manner, blocking is partially responsible for setting the water properties in the eastern basins. At shallower depth, a broad-scale slope of isotherms downwards to the east is seen, denoting existence of an eastward baroclinic pressure gradient that drives the less-dense bottom water along the RFZ. Three localized regions are characterized by particularly abrupt descent of isotherms. These correspond to topographical features within the RFZ. The first is a horizontal constriction at $0^\circ 47'\text{N}$, $13^\circ 25'\text{W}$, the second a sill immediately downstream from the constriction (sill 1, at $0^\circ 47'\text{N}$, $13^\circ 18'\text{W}$). The third consists of a broad sill (sill 2) followed by a sharp increase in bottom depth at $0^\circ 55'\text{N}$, 13°W . Dramatic changes in the flow can be inferred from the sloping isotherms and the principle of energy conservation. The bottom water accelerates downstream at these sites, with potential energy exchanged for kinetic energy along the channel. The velocity profile data acquired by the HRP (Fig. 2b) support this idea; bottom currents increase from $\sim 25\text{ cm s}^{-1}$ at the main sill to near 50 cm s^{-1} . Immediately downstream from these topographical features, deep isotherms abruptly shoal on still-smaller horizontal scales, suggesting the possibility of internal hydraulic jumps.

The flow accelerations provide for extremely large vertical shear within the AABW flow. The Richardson number¹⁴ associated with the large-scale flow (estimated over a 100-m vertical scale) was less than 0.25 for most vertical profiles downstream of the main sill, indicating supercriticality to shear instability. We believe that the substantial vertical gradients of horizontal velocity support shear instability¹⁵ within the bottom-water flow and in turn, turbulent mixing (see below). Hydraulic theory suggests an additional explanation for the observed turbulence. Estimates of Froude number based on two-layer hydraulic theory¹⁶ indicate that the flow of bottom water is hydraulically controlled just downstream of the two small sills. Hydraulic control implies that the height of bottom water over the sill uniquely determines the velocity, transport and height of bottom water upstream of the sill,

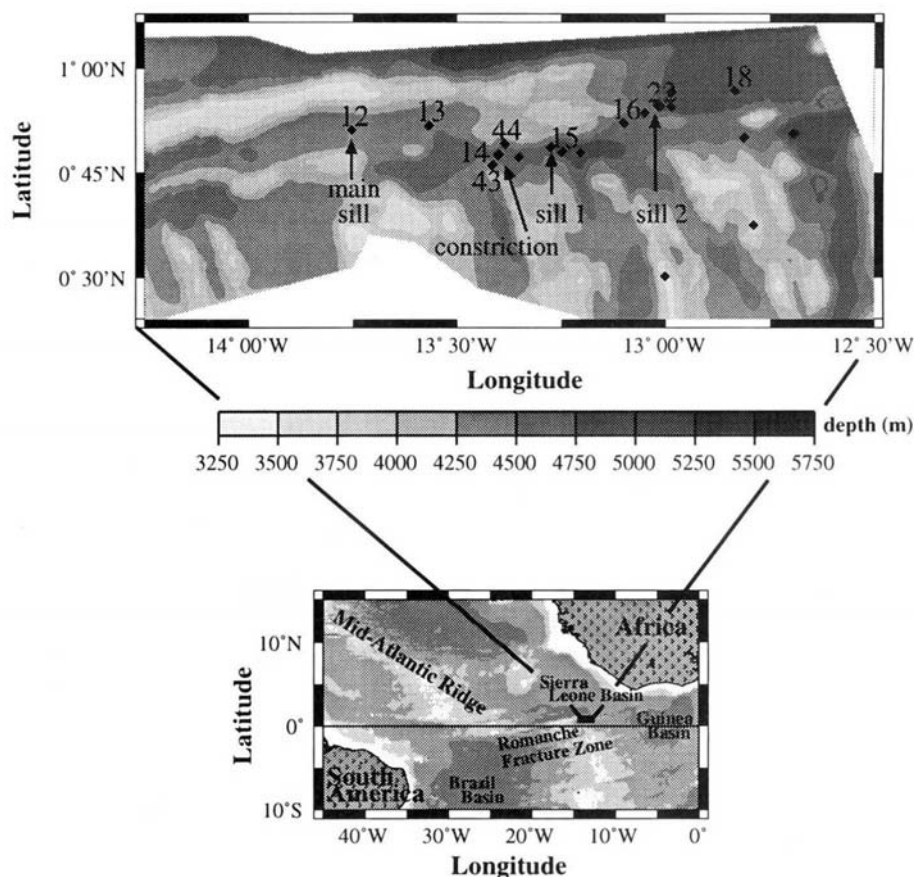


FIG. 1 Bathymetry of the Romanche fracture zone from multi-beam depth sounder data (H. Mercier, personal communication), with positions of selected HRP stations discussed in the text. Light shading denotes shallow topography. This fracture zone represents a principal conduit for Antarctic Bottom Water flow from the western basins of the Atlantic to those in the east.

similar to flow over a weir. Hydraulic jumps, denoting the transition from a controlled to a sub-critical state downstream, are often sites of intense turbulent mixing¹⁷.

Exceptionally energetic turbulent velocity fluctuations were observed east of the main sill, particularly downstream of the topographical features discussed previously (Fig. 2c). The intensity of the turbulence was quantified with estimates of the turbulent kinetic energy dissipation rate¹⁸. Maximum dissipation rates, in excess of $1 \times 10^{-6}\text{ W kg}^{-1}$, were observed just past the second sill (site of dive 22), where consistently large values of dissipation were found in a series of repeated profiles. The heightened levels of turbulence were generally confined to the bottom water. This was not exclusively the case, however, as dive 22 at the second sill made evident. At this sill, the layer of vigorous turbulence extended well above the 1.9°C isotherm, indicating mixing between bottom water and overlying North Atlantic Deep Water.

Looking east along the flow direction past the main sill, the coldest isotherms are seen successively to disappear in Fig. 2a, providing qualitative evidence of mixing processes². We estimated the turbulent eddy diffusivity from the kinetic energy dissipation data for the stations shown in Fig. 2c (ref. 19). For the waters below 4,000 m within the RFZ, downstream from the main sill, the average diapycnal diffusivity was $150 \times 10^{-4}\text{ m}^2\text{ s}^{-1}$. This contrasts with the estimate of the mid-depth diffusivity at these and other stations in the region of $0.2 \times 10^{-4}\text{ m}^2\text{ s}^{-1}$. Our sampling was spatially coarse; it is possible that we may have missed spots with even larger turbulent intensities in this preliminary survey. Estimates of water mass conversion rates based on large-scale velocity and temperature observations will be made for comparison with the microstructure measurements in future analysis. Nevertheless, the velocity estimates coupled with the levels of

turbulence we did observe makes obvious the conclusion that turbulent mixing contributes significantly to the downstream evolution of the temperature field.

A cross-channel section consisting of three profiles (14, 43 and 44) was occupied just downstream of the constriction (13° 25' W, Fig. 1) in order to estimate the transport of bottom water. At this site, the average depth of the 1.9 °C potential-temperature surface (defining the top of the bottom water) is ~3,940 m. The channel on this section has a maximum depth of 4,707 m. The northern channel wall along the section line rises to ~4,050 m, reaches a plateau, then rises again to depths shallower than 4,000 m. To the south, the channel wall is formed by a ridge which extends in the southeast direction. Along the section line the bathymetry arises to only 4,200 m at the ridge crest, but following the ridge axis southeast the bottom shoals to well under 4,000-m depth. Our sampling pattern thus delineates the flow of the densest waters through the RFZ (those waters below ~4,200 m) but the upper portion of the bottom water flow is not as well constrained by these observations. Confining the integration of the velocity profiles in the horizontal by the 4,050-m bathymetric contour to the north and the 4,200-m contour to the south, we obtain a transport of $1.0 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ assuming a zero-velocity surface at the 2.1 °C potential-temperature surface (where the large-scale temperature field suggests the zonal pressure gradient is near zero). By extrapolation, we estimate possible additional contributions of $0.2 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ to the north and $0.1 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ to the south. Our transport estimate can be compared to previous indirect estimates of $< 1.2 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ (ref. 20), $2 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ (ref. 21) and $(2.6\text{--}5.1) \times 10^6 \text{ m}^3 \text{ s}^{-1}$ (ref. 22). Although flow along the RFZ may exit southwards into the region of broken topography before it reaches the HRP section line, it has been suggested that this flow is small and directed northwards, not southwards². Our measurements in the deep valleys between the Romanche and Chain fracture zones support the inference of weak flow there. In addition, from our two profiles in the Chain fracture zone, we estimate the bottom-water transport there to be only $\sim 0.1 \times 10^6 \text{ m}^3 \text{ s}^{-1}$. The present measurements suggest a relatively small total transport of AABW through the Romanche and Chain fracture zones of $\sim 1.4 \times 10^6 \text{ m}^3 \text{ s}^{-1}$.

Current-meter and CTD measurements of dense-water transport through bathymetric constrictions such as the RFZ have been used previously to infer average mixing rates^{23–27}. These estimates are deduced from heat budgets for bottom waters in semi-enclosed basins downstream from constrictions, and are frequently expressed as an average diffusivity, K_a . The estimates of K_a typically range between 1 and $10 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$. Importantly, these studies do not clearly differentiate between modest-intensity, spatially homogeneous mixing in the basin interior and intense mixing confined to the sill region. We have presented evidence that exceptionally high levels of turbulent dissipation supported by shear instability can occur in the abyssal ocean within narrow passages characterized by intense flow. In the context of the RFZ, the estimated diffusivity of $150 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ acting over the $\sim 10^3 \text{ km}^2$ area

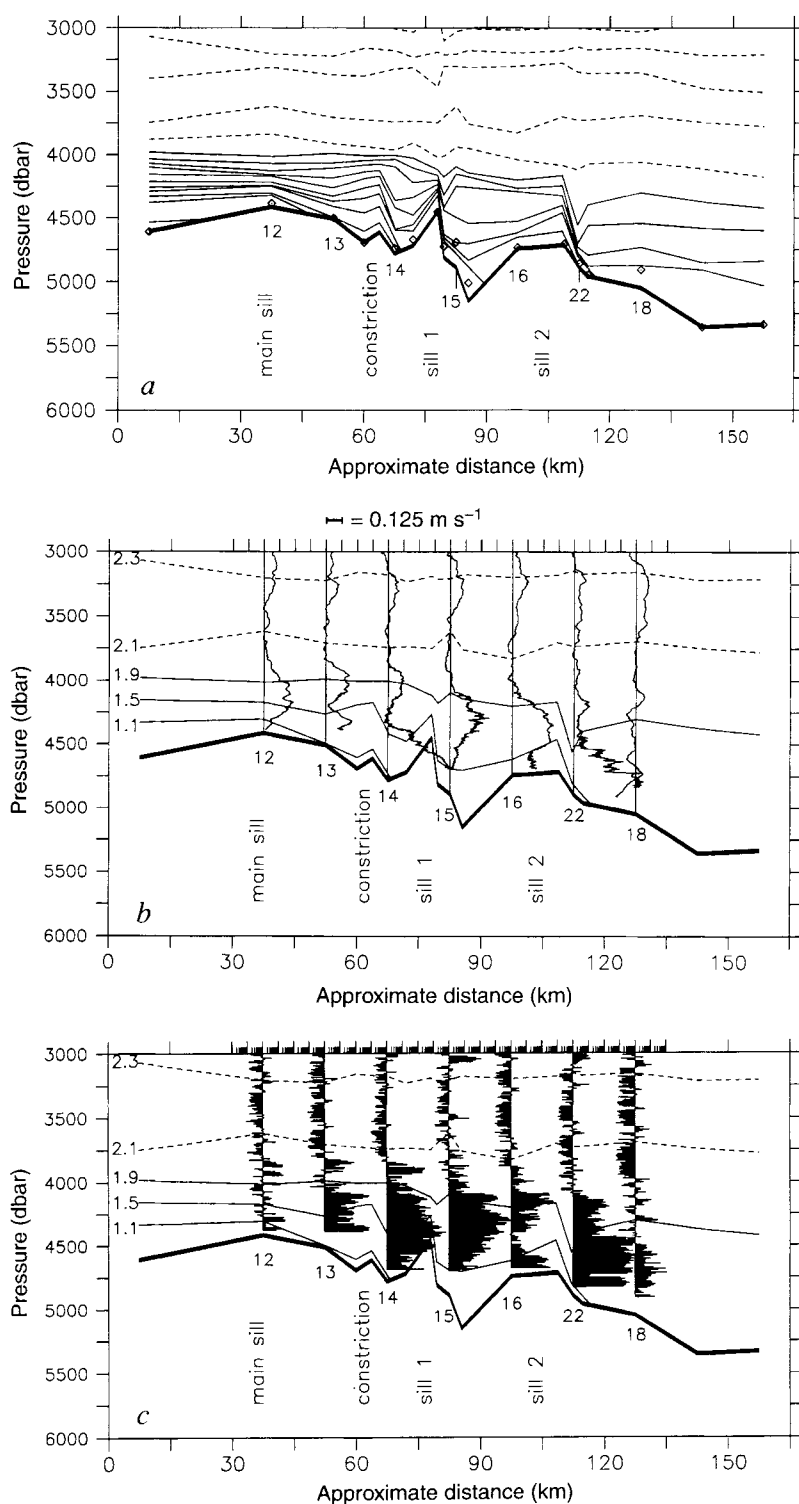


FIG. 2 Along-channel sections of potential temperature, velocity and turbulent dissipation. To enhance clarity, the separation between stations (12, 13, 14 and so on) is approximate. The section spans 160 km, and the bold line denotes bottom topography (pressure in dbar $\approx$ depth in m). a, Potential temperature. The contour interval is 0.1 °C. Solid contours represent temperatures ≤ 1.9 °C (Antarctic Bottom Water). The symbols correspond to the greatest pressure attained by either the HRP or a contemporaneous CTD profile. Additional data from CTD work completed in 1991 (H. Mercier, personal communication) are included to refine the picture. The earlier data are virtually identical to the present data in the case of co-located measurements (for example, bottom potential temperatures differ by < 0.02 °C). b, Along-channel velocity. The velocity profiles at each station were rotated into along- and across-channel coordinates and referenced by setting the velocity at the depth of the 2.1 °C potential-temperature isotherm to zero. The velocity scale appears above the figure. Some potential temperature isotherms (labelled 1.1, 1.5, 1.9, 2.1 and 2.3; all °C) are also shown. c, Turbulent dissipation rate. Profiles of this quantity are displayed at each station; note the logarithmic axis. The data are shown with a reference value of $10^{-10} \text{ W kg}^{-1}$, so that values $> 10^{-10}$ appear as horizontal lines to the right of the vertical. Potential-temperature contours are also shown as in b.

downstream of the main sill is equivalent to an interior diffusivity of $0.1 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ acting over an area of $1.5 \times 10^6 \text{ km}^2$, a region similar in areal extent to the Sierra Leone Basin.

It is not clear to what degree these observations of intense mixing within the Romanche fracture zone can be generalized to other deep sill flows. Although one might expect hydraulically controlled flows to exhibit intense mixing, as has been demonstrated in the Strait of Gibraltar for example²⁸, it has not been established that the bottom waters for which abyssal heat budgets have been done are controlled in the same manner. Another issue to be resolved is the relative importance of interior versus near-bottom mixing away from flow constrictions. As only a few exploratory measurements of abyssal mixing rates have yet been obtained, additional observations identifying the sites and processes of deep-water mass conversion will be essential before the global thermohaline budget can be closed. As evidenced by the dramatic warming of AABW crossing the Romanche fracture zone, localized mixing exerts a profound influence on the water property characteristics of the eastern basins of the North Atlantic. Accurate parametrization of these mixing processes is a vital step in developing realistic models of ocean circulation. □

Received 29 August 1995; accepted 18 January 1996.

1. Wüst, G. Das Bodenwasser und die Gliederung der Atlantischen Tiefsee. in *Wissenschaftliche Ergebnisse der Deutschen Atlantischen Expedition auf dem Forschungs- und Vermessungsschiff "Meteor" 1925–1927*, Vol. 6(1), 1–107 (Gruyter, Berlin, 1933).

2. Mercier, H., Speer, K. G. & Honnorez, J. *Deep-Sea Res.* **41**, 1457–1477 (1994).
3. Garrett, C. & Munk, W. *Deep-Sea Res.* **19**, 823–832 (1972).
4. Garrett, A. E. & Holloway, G. *J. mar. Res.* **42**, 15–27 (1984).
5. Munk, W. H. *Deep-Sea Res.* **13**, 707–730 (1966).
6. Armi, L. J. *geophys. Res.* **83**, 1971–1979 (1978).
7. Eriksen, C. C. J. *phys. Oceanogr.* **15**, 1145–1156 (1985).
8. Bell, T. H. J. *geophys. Res.* **80**, 320–327 (1975).
9. Toole, J. M., Polzin, K. L. & Schmitt, R. W. *Science* **264**, 1120–1123 (1994).
10. Kunze, E. & Sanford, T. M. J. *phys. Oceanogr.* (submitted).
11. Schmitt, R. W., Toole, J. M., Koehler, R. L., Doherty, K. & Mellinger, E. J. *Atmos. Oceanic Technol.* **5**, 484–500 (1988).
12. Monti, S. & Mercier, H. *Bathymetric Map of the Romanche Fracture Zone* (IFREMER, Brest, France, 1991).
13. Bryden, H. L. & Stommel, H. M. *Oceanologica Acta* **7**, 289–296 (1984).
14. Miles, J. W. & Howard, L. N. J. *Fluid Mech.* **10**, 496–512 (1961).
15. Thorpe, S. A. J. *Fluid Mech.* **61**, 731–751 (1973).
16. Armi, L. J. *Fluid Mech.* **163**, 27–58 (1986).
17. Long, R. R. *Bull. Am. met. Soc.* **34**, 205–211 (1953).
18. Tennekes, H. & Lumley, J. L. *A First Course in Turbulence* (MIT Press, Cambridge, 1972).
19. Osborn, T. R. J. *phys. Oceanogr.* **10**, 83–89 (1980).
20. McCartney, M. S. & Curry, R. A. J. *phys. Oceanogr.* **23**, 1264–1276 (1993).
21. Warren, B. A. & Speer, K. G. *Deep-Sea Res.* (suppl.) **38**, s281–s322 (1991).
22. Schlitzer, R. J. *geophys. Res.* **92**, 2953–2969 (1987).
23. Whitehead, J. A. & Worthington, L. V. J. *geophys. Res.* **87**, 7903–7924 (1982).
24. Hogg, N., Biscaye, P., Grdner, W. & Schmitz, W. J. *mar. Res.* (suppl.) **40**, 231–263 (1982).
25. Saunders, P. M. J. *phys. Oceanogr.* **17**, 631–643 (1987).
26. Barton, E. D. & Hill, A. E. *Deep-Sea Res.* **36**, 1121–11226 (1989).
27. Roemmich, D., Hautala, S. & Rudnick, D. J. *geophys. Res.* (in the press).
28. Wessen, J. C. & Gregg, M. C. J. *geophys. Res.* **99**, 9847–9878 (1994).

ACKNOWLEDGEMENTS. We thank the following for their essential contributions; E. Montgomery, D. Wellwood and T. Bolmer to HRP operations, J.-P. Girardot and T. Penduff for CTD work, N. Oakley for shear probe calibration assistance and advice, and Le Noroit Commandant G. Trédunit and his crew for ship handling. We also thank H. Mercier and H. Bryden for stimulating our research. This work was supported by the US NSF and the French CNRS, IFREMER and UBO.

A 4-Gyr shock age for a martian meteorite and implications for the cratering history of Mars

R.D. Ash^{*†}, S. F. Knott^{*} & G. Turner^{*}

^{*} Department of Earth Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

[†] Department of Earth and Planetary Sciences, American Museum of Natural History, Central Park West at 79th Street, New York 10024-5192, USA

ANALYSES of meteorites that originated on Mars provide important insights into the geological and atmospheric evolution of the planet. Such analyses have hitherto been restricted to relatively young martian rocks¹ (the oldest martian meteorites have an age of approximately 1.3 billion years). But the recently recognized² martian meteorite, Allan Hills 84001, which is distinct from the other martian meteorites^{2–4}, shows evidence for a much older age^{5,6}. Here we report an analysis of the shock-alteration history of this meteorite based on argon isotope dating, from which we derive a shock age of 4.0 ± 0.1 billion years. The age and geological history of this meteorite suggest that it came from the heavily cratered Noachian-age terrains of Mars's southern hemisphere, and it may thus provide an absolute chronology for this region of the planet, independent of that inferred from the cratering record. The shock age of the meteorite also coincides with that of the so-called Lunar Cataclysm (a relatively short period during which many of the craters on the Moon are believed to have formed), supporting the idea⁷ that intense bombardment was widespread throughout the inner Solar System between 3.9 and 4.1 billion years ago.

The chronology of Mars is based on stratigraphic relationships between surface features, with absolute age estimates of martian stratigraphic units being made by comparing its crater densities with those of the Moon⁸. But these ages are severely model-dependent and rely on the extrapolation of the ancient cratering record of the Moon to that of Mars. The oldest, most heavily

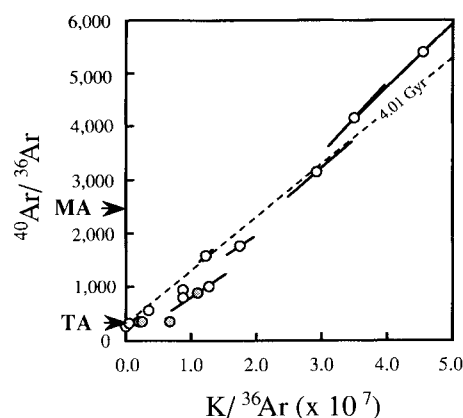


FIG. 1 Plot of $^{40}\text{Ar}/^{36}\text{Ar}$ against $\text{K}/^{36}\text{Ar}$ of ALH84001.110 showing the correlation between K and ^{40}Ar . Points with excess ^{39}Ar are low-temperature terrestrial K-rich salts (filled symbols) and high-temperature ^{39}Ar recoil implanted pyroxene (see Fig. 2 legend), and correspond to the lower ages in the stepped release profile. Martian atmosphere (MA) $^{40}\text{Ar}/^{36}\text{Ar}$ ratio is $2,400 \pm 100$ (ref. 21), terrestrial atmospheric (TA) ratio is 295.5. Data are all corrected for cosmogenic ^{36}Ar . The 4.01-Gyr isochron is drawn for reference purpose only with a slope corresponding to the total gas released above 400°C .

cratered regions of Mars lie in the southern hemisphere and have been termed Noachian-age terrains with age estimates ranging from 3.0 Gyr (ref. 9) to 4.4 Gyr (ref. 10). The northern hemisphere is generally less heavily cratered with ages for the youngest features, such as Olympus Mons, estimated between 0.1 Gyr (ref. 9) to 2.5 Gyr (ref. 10). Furthermore, dating of lunar samples indicates that the Moon underwent a period of intense bombardment—the Lunar Cataclysm—over a brief period between 3.9 and 4.0 Gyr ago^{11,12} but whether this was a local phenomenon or was also experienced by Mars remains unresolved^{13,14}.

Mineralogical, chemical and isotopic analyses confirm a close link between ALH84001 and the SNC meteorites²—a group of meteorites for which there is overwhelming evidence for a martian

origin¹—thus establishing ALH84001 as a rock from Mars. But ALH84001 differs from the other martian meteorites, exhibiting a complex geological history² consisting of at least two shock deformation episodes, thermal metamorphism³ and aqueous alteration⁴. This had led to the suggestion that this meteorite is older than the other martian meteorites, most of which have ages of 1.3 Gyr or younger¹. Results from Sm–Nd and Rb–Sr dating of ALH84001 acid leaches and mineral separates support this argument, suggesting igneous crystallization ages of ~4.56 Gyr (refs 5, 6). Whilst these age data provide information about igneous differentiation and crystallization processes, meteorite shock histories are more open to examination by the Ar–Ar dating technique¹⁵. We have applied the latter method in the hope of obtaining more information about the evolution of ALH84001 and its parent-planet cratering history.

Previous analyses of argon extracted from ALH84001 have been carried out on unirradiated samples, and age estimates have been relatively unconstrained owing to the unknown contributions from martian and terrestrial atmosphere and difficulties in determining the potassium content of the samples analysed. The advantages offered by Ar–Ar dating are that the multiple contributors to the overall Ar budget (martian and terrestrial atmosphere, cosmogenic and radiogenic Ar and reactor-produced isotopes from neutron reactions) can be deconvoluted and the potassium content determined directly on the samples analysed from the conversion of ³⁹K to ³⁹Ar, which is then released on stepped pyrolysis. We have extracted argon from three separate irradiated aliquots of ALH84001 by incremental heating. The bulk argon contents and compositions are similar to those found in previous work^{16,17} and the determinations of bulk potassium content, between 64 and 101 parts per million (p.p.m.), are in reasonable agreement with a previously determined value¹⁸ of 108 ± 16 p.p.m. Microprobe analysis of the samples show that maskelynite—a shock-produced feldspathic glass—is the major potassium-bearing phase, containing up to 2.4 wt% K₂O.

Figure 1 shows a ³⁶Ar-normalized plot of the ⁴⁰Ar content against the potassium content (from the release of ³⁹Ar) for each of the temperature increments of the stepped heating of one of the samples (ALH84001.110) which are clearly correlated. The Ar release is consistent with a two-component mixture of

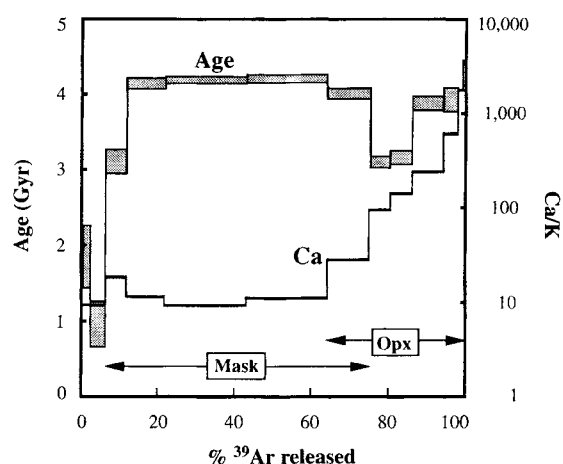


FIG. 2 The stepped heating of ALH84001.110. The cumulative K release is determined from the ³⁹Ar release, and age determinations are corrected for Ca and Cl interferences, cosmogenic Ar production and terrestrial atmospheric blanks (see ref. 22 for full correction procedure). The drop in apparent age at high temperature is a common feature in meteorite Ar–Ar age profiles and is generally attributed to the recoil of ³⁹Ar, during the reactor neutron irradiation, from K-rich phases (maskelynite) into more refractory, K-poor phases (orthopyroxene). 'Mask' and 'Opx' refer to the dominant release temperatures of maskelynite and orthopyroxene respectively.

TABLE 1 Results of analysis of three samples of ALH84001

Sample number	0.110	0.111	0.127
Weight (mg)	21.0	29.2	14.0
K (p.p.m.)	101	64	91
⁴⁰ Ar*/K ($\times 10^{-6}$)	100.7 ± 0.9	98.4 ± 4.1	107.1 ± 5.7
Age (Gyr)	$4,010 \pm 14$	$3,940 \pm 70$	$4,110 \pm 90$

These stepped heating analyses were performed on three samples of ALH84001. Sample ALH84001.110, unaltered interior chip; *ALH84001.111, altered interior chip; ALH84001.127, carbonate-rich altered interior sample. ⁴⁰Ar* is radiogenic ⁴⁰Ar from the *in situ* decay of ⁴⁰K.

radiogenic ⁴⁰Ar and a trapped terrestrial-like component. But we find little evidence for the presence of a significant martian atmospheric component, which would result in the data plotting within a triangle defined by the terrestrial and martian atmospheres and a radiogenic component, the composition of which would be defined by the potassium content and age. The same data is plotted as age versus cumulative release of ³⁹Ar in Fig. 2 but the analyses of all the ALH84001 samples follow an essentially identical pattern. Each of the samples show low apparent ages at low and high temperatures. The former are probably due to the partial degassing of ⁴⁰Ar and contamination by K-rich terrestrial salts, the latter due to ³⁹Ar recoil (see Fig. 2 legend). For these reasons, the best overall age is given by the total Ar release above 400 °C. Using this constraint, the ages range from 3.9 to 4.1 Gyr compared with 4.1–4.2 Gyr using a simple plateau age. The results from the stepped-heating analyses of the three samples are given in Table 1.

The petrography of ALH84001 shows that, after cooling from an igneous melt, it was subject to at least one shock event, followed by metamorphic heating and then chemical alteration during which carbonates were formed. There then followed a further shock event which produced the observed melting and led to maskelynite formation³. As the maskelynite contains >96% of the total potassium, it is clearly the degassing of this phase which is being dated. However owing to the complexity of events experienced by ALH84001, it is not clear exactly which of these events led to the degassing—though it is unlikely that it is this event that caused the meteorite ejection from Mars as the survival of small stony bodies in space is comparatively short. The timing of that particular event is more likely to be indicated by the 15-Myr cosmic-ray exposure age^{16,17}. It seems more probable that the 4.0-Gyr age reflects argon loss caused by an earlier cratering impact on the surface of Mars. The degassing during this event may have been due to the production of maskelynite or to a period of thermal annealing (metamorphism), possibly in an ejecta blanket. Thus there is evidence that this rock has experienced at least two, probably three, substantial impact events and that it has an old age—hence this rock is likely to have been derived from an ancient, highly cratered surface which, on Mars, are the Noachian-age terrains in the south.

A heavy shock age of 3.9–4.1 Gyr for this old martian sample coincides with that of the Lunar Cataclysm or Late Heavy Bombardment of the Moon. It is also an Ar–Ar age which has been measured for a group of achondritic (howardite–eucrite–diogenite association) meteorites⁷ for which there is spectroscopic evidence^{19,20} for their genesis on the asteroid Vesta 4. Furthermore it has been noted that there is a peak in the ages of shocked ordinary chondrites at this time⁷ (see Fig. 3). From these age data it has been argued that there was a bombardment throughout the inner Solar System at a relatively late time, commencing some 4.0 Gyr ago⁷. Whilst caution must clearly be exercised in drawing conclusions from a single rock sample, it is suggestive that the shock age determined for the first old martian meteorite coincides with this period of late bombardment. It may be interpreted as the first evidence that Mars too was subject to such a bombardment,

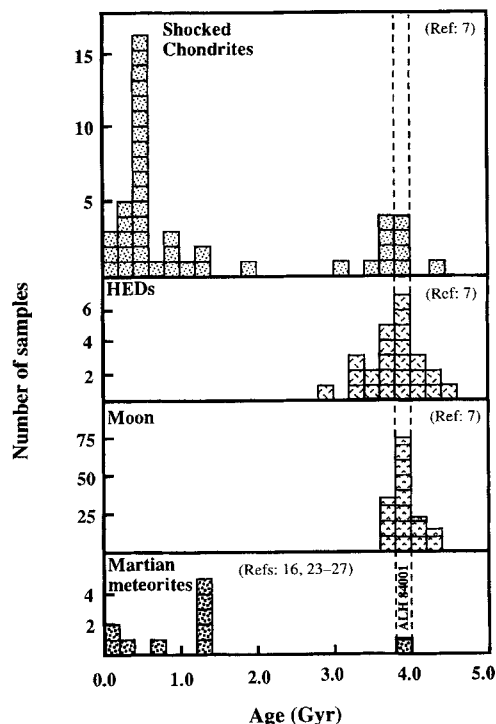


FIG. 3. Ar-Ar shock ages for meteorites and lunar highland samples showing the coincidence of shock ages at ~4 Gyr. Each plotted time interval is 200 Myr. The HED (howardite-eucrite-diogenite) ages used are those referred to as "precise"¹⁷.

and that the ancient Noachian-age terrains are from this era. But the full significance of this age and further insights will only be gleaned on the finding and analysis of future old Martian meteorites or, ultimately, the analysis of samples from return missions to Mars and other bodies in the inner Solar System. □

Received 25 July 1995; accepted 29 January 1996.

1. McSween, H. Y. Jr *Meteoritics* **29**, 757-779 (1994).
2. Mittlefeldt, D. W. *Meteoritics* **29**, 214-221 (1994).
3. Trieman, A. H. *Meteoritics* **30**, 294-302 (1995).
4. Romanek, C. S. et al. *Nature* **372**, 655-657 (1994).
5. Jagoutz, E., Sorrow, A., Vogel, J. D. & Wänke, H. *Meteoritics* **29**, 478-479 (1994).
6. Nyquist, L. E., Bansal, B. M., Wiesmann, H. & Shih, C.-Y. *Lunar planet Sci.* **XXVI**, 1065-1066 (1995).
7. Bogard, D. D. *Meteoritics* **30**, 244-268 (1995).
8. Neukum, G. & Hiller, K. J. *geophys. Res.* **86**, 3097-3121 (1981).
9. Hartmann, W. K. J. *geophys. Res.* **78**, 4096-4116 (1973).
10. Neukum, G. & Wise, D. U. *Science* **194**, 1381-1387 (1976).
11. Tera, F., Papanastassiou, D. A. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **22**, 1-21 (1974).
12. Turner, G. *Phys. Chem. Earth* **10**, 145-195 (1977).
13. Ryder, G. *Lunar planet. Sci.* **XX**, 934-935 (1989).
14. Strom, R. G., Croft, S. K. & Barlow, N. G. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 383-423 (Univ. Arizona Press, Tucson, 1992).
15. Turner, G. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 276-288 (Univ. Arizona Press, Tucson, 1988).
16. Miura, Y. N., Nagao, K., Sugiyama, N., Sagawa, H. & Matsubara, K. *Geochim. cosmochim. Acta* **59**, 2105-2114 (1995).
17. Swindle, T. D., Grier, J. A. & Burkland, M. K. *Geochim. cosmochim. Acta* **59**, 793-801 (1995).
18. Mittlefeldt, D. W. *Meteoritics* **29**, 900 (1994).
19. Drake, M. J. in *Asteroids* (ed. Gehrels, T.) 765-782 (Univ. Arizona Press, Tucson, 1988).
20. Binzel, R. P. *Meteoritics* **30**, 486-487 (1995).
21. Swindle, T. D., Caffee, M. W. & Hohenberg, C. M. *Geochim. cosmochim. Acta* **50**, 1001-1025 (1986).
22. Knott, S. F., Ash, R. D. & Turner, G. *Lunar planet. Sci.* **XXVI**, 765-766 (1995).
23. Bogard, D. D., Husain, L. & Nyquist, L. E. *Geochim. cosmochim. Acta* **43**, 1047-1055 (1979).
24. Shih, C.-Y. et al. *Geochim. cosmochim. Acta* **46**, 2323-2344 (1982).
25. Podosek, F. A. *Earth planet. Sci. Lett.* **19**, 135-144 (1973).
26. Bogard, D. D. & Husain, L. *Geophys. Res. Lett.* **4**, 49-71 (1977).
27. Bogard, D. D. & Nyquist, L. E. *Meteoritics* **14**, 356 (1979).

ACKNOWLEDGEMENTS. We thank R. Burgess for his help and technical expertise, J. D. Gilmour for his elliptical error plotting capabilities, and the Antarctic Meteorite Working Group for providing samples. We also thank C. Chapman for helpful comments and suggestions. This work was supported by the NERC.

Detection of magnetic field intensity by sea turtles

Kenneth J. Lohmann & Catherine M. F. Lohmann

Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA

WHETHER migratory animals can determine their global position by detecting features of the Earth's magnetic field has long been debated¹⁻⁴. To do this an animal must perceive (at least) two distinct magnetic parameters, each of which must vary in a different direction across the Earth's surface^{3,5}. There has been no evidence that any animal can perceive two such magnetic features, and whether 'magnetic maps' exist at all has remained controversial²⁻⁶. Several populations of sea turtles⁷⁻⁹ undergo transoceanic migrations before returning to nest on or near the same beaches where they themselves hatched. Along the migratory routes, all or most locations have unique combinations of magnetic field intensity and field line inclination. It has been demonstrated that hatchling loggerhead turtles can distinguish between different magnetic inclination angles¹⁰. Here we report that turtles can also distinguish between different field intensities found along their migratory route. Thus sea turtles possess the minimal sensory abilities necessary to approximate global position using a bicoordinate magnetic map.

The ability to return to a specific beach for nesting from hundreds or thousands of kilometres away is common among sea turtles. Green turtles (*Chelonia mydas*) that nest on Ascension Island in the South Atlantic regularly migrate between their nesting beach and their feeding grounds in Brazil, a straight-line distance of more than 2,000 km (refs 11, 12). Most Kemp's ridley turtles (*Lepidochelys kempi*) throughout the Atlantic, Caribbean and the Gulf of Mexico converge on a single, isolated beach in Mexico to lay their eggs^{11,13}, and loggerheads (*Caretta caretta*) that nest in Japan cross the Pacific Ocean to Baja California before returning to their natal beaches to nest⁹.

Adult sea turtles can also pinpoint small, isolated sites that are not nesting beaches. Loggerheads and green turtles that nest along the Great Barrier Reef return to specific, widely dispersed feeding grounds located hundreds or thousands of kilometres away from their nesting sites¹⁴. Recent satellite tracking experiments have shown that migrating green turtles often follow essentially straight paths directly to their goals¹⁵. How sea turtles



FIG. 1 Diagram of the North Atlantic gyre. Currents of the gyre are represented by arrows. The number 1 indicates the only location within the gyre where the field intensity is 52,000 nT; 2 indicates where Florida loggerheads in the gyre presumably first encounter a field intensity of 43,000 nT.

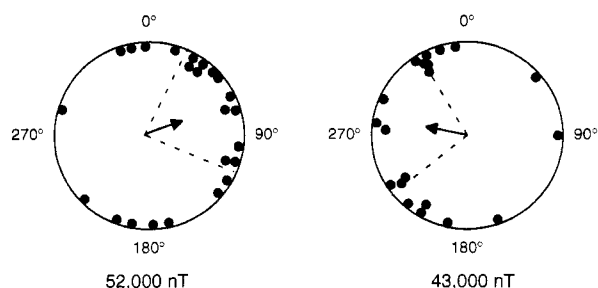


FIG. 2 Orientation of hatchling loggerheads tested in a magnetic field of 52,000 nT (left) and 43,000 nT (right). The inclination angle in each case matched that found at the turtles' natal beach in Florida (57°). Each dot represents the mean angle of a single hatchling. The arrow in the centre of each circle indicates the mean angle of the group; the arrow length is proportional to the magnitude of the mean vector r , with the radius of the circle corresponding to $r = 1$. Turtles tested in the field of 52,000 nT were significantly oriented ($r = 0.40$, $Z = 3.75$, $P = 0.020$, Rayleigh test) with a mean angle of 69° . Turtles tested in the field of 43,000 nT were also significantly oriented ($r = 0.41$, $Z = 3.29$, $P = 0.033$), but in approximately the opposite direction (mean angle, 280°). Broken lines indicate the 95% confidence interval for the mean angle.

METHODS. Methods have been described in detail previously¹⁰. Each turtle was tethered to an electronic tracking unit in the centre of a water-filled orientation arena. This was surrounded by a computerized coil system, which was used to control the magnetic field intensity and inclination. Each turtle began its trial in a magnetic field matching that found at the natal beach (inclination angle 57° , intensity 47,000 nT) and was allowed to establish a course towards a dim light in magnetic east¹⁰. The light was then turned off, and the field intensity was simultaneously either increased to 52,000 nT or decreased to 43,000 nT, with the inclination angle remaining unchanged. After an acclimation period of 3 min, a computer monitored the direction that each turtle swam towards in darkness under the new field condition¹⁰.

accomplish these remarkable feats of long-distance navigation is not known.

Like adults, hatchling sea turtles are migratory. Loggerhead hatchlings along the east coast of Florida swim offshore immediately after emerging from their nests¹⁶. They eventually reach the North Atlantic gyre¹⁷ (the circular current system that encircles the Sargasso Sea; see Fig. 1), and remain within the confines of the gyre for several years as they grow and mature^{17,18}.

Because hatchlings are small and more amenable to laboratory experiments than adults, research on sea-turtle orientation and navigation has focused largely on hatchlings. Hatchling loggerheads can orient to the Earth's magnetic field^{19,20}, and turtles can distinguish between magnetic inclination angles found at different points along their migratory route in the North Atlantic gyre¹⁰.

Magnetic field intensity, like inclination, varies across the surface of the Earth. To determine whether hatchlings can discriminate between different field intensities, we subjected hatchling loggerheads to two different intensities found along their migratory route but on opposite sides of the Atlantic Ocean (Fig. 1). Turtles were tested in a circular, water-filled arena that was surrounded by a computerized coil system¹⁰, which was used to control precisely the magnetic field in which each turtle swam. Each hatchling was tethered to an electronic tracking unit that relayed the position of the turtle to a computer in an adjacent room¹⁰.

Turtles tested in a field of 52,000 nT (a field 10.6% stronger than the natal beach field, and at a level only encountered by hatchlings near South and North Carolina, U.S.A.) swam approximately eastwards (Fig. 2). Those exposed to a field of 43,000 nT (a field 8.5% weaker than the natal beach field, and at a level first encountered on the eastern side of the Atlantic, near Portugal) swam westwards (Fig. 2). The two distributions are significantly different (Watson test²¹, $U^2 = 0.401$, $P < 0.001$).

These results demonstrate that hatchlings can distinguish

between magnetic field intensities that are normally encountered during their first months in the North Atlantic gyre. In addition, the directional responses are consistent with the hypothesis that hatchlings derive positional information from field intensity, and then swim in directions that keep them within the gyre's boundaries.

Within the gyre, a field intensity of 52,000 nT exists at only one location: along the southeastern United States coast, in the region where the Gulf Stream veers eastwards across the Atlantic (Fig. 1). Thus the eastward orientation of hatchlings in a field of 52,000 nT is consistent with the migratory route of Florida loggerheads^{17,18}. Such orientation would presumably lead turtles farther from the North American coast and towards the centre of the Gulf Stream, which in turn might reduce the likelihood of turtles straying into fatally cold water north of the gyre^{10,18}. Similarly, because turtles first encounter a field of 43,000 nT on the eastern side of the Atlantic, the westward orientation observed under this condition might also function to keep turtles within the confines of the gyre.

Although the results demonstrate that hatchlings can derive at least some positional information from features of the Earth's magnetic field, whether hatchlings can actually fix their position relative to a goal (that is, whether they can navigate as adults do) is not known. Conceivably, hatchlings might emerge from their nests programmed merely to swim in specific directions in response to particular magnetic features found along the migratory route (for example, magnetic parameters along the far northern border of the gyre might elicit southward orientation, whereas features along the southern border of the gyre might elicit northward swimming)¹⁰. Thus young turtles might remain within a favourable gyre or other oceanic region without possessing an ability to navigate.

In contrast, the ability of adults to swim directly to small target sites hundreds or thousands of kilometres away clearly implies an ability to fix position relative to a goal^{12,15,16}. Given that hatchlings can detect both magnetic inclination angle and magnetic field intensity, we hypothesize that older sea turtles learn the gradients of these two magnetic features and eventually, as adults, develop large-scale, bicoordinate magnetic maps for use in long-distance migrations.

A particularly clear example of how inclination and intensity might be used in position finding can be seen in the South Atlantic, over the entire migratory range of the Ascension Island green turtle population (Fig. 3). Between Ascension and Brazil, isoclinics (lines of equal field inclination) and isodynamics (lines of equal field intensity) form a nearly orthogonal grid. Thus all locations between the feeding grounds in Brazil and the nesting beaches at Ascension are defined by unique combinations of

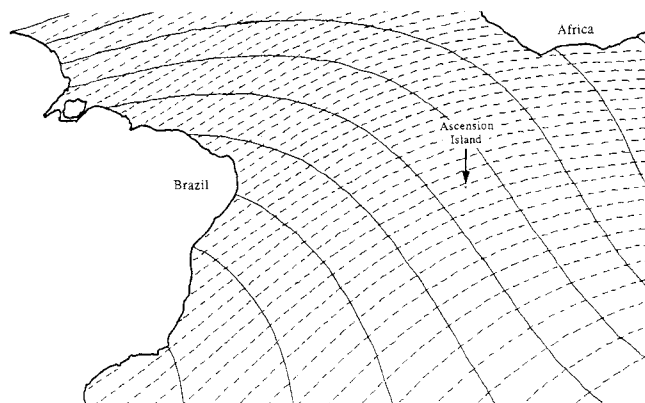


FIG. 3 Isoclinics²⁷ (broken lines) and isodynamics²⁸ (solid lines) in the oceanic region surrounding Ascension Island. Adjacent isoclinics differ by 2° and adjacent isodynamics by 1,000 nT. The two geomagnetic features form a non-orthogonal grid that might, in principle, provide Ascension Island turtles with a bicoordinate position-finding system as they migrate between Ascension and the Brazilian coast.

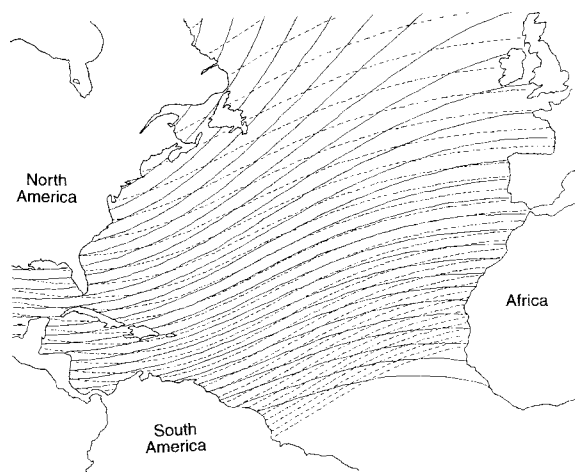


FIG. 4 Isoclinics²⁷ (broken lines) and isodynamics²⁸ (solid lines) in the North Atlantic. For conventions, see Fig. 3.

intensity and inclination. A migrating turtle using a bicoordinate magnetic map based on these two parameters could therefore determine its position anywhere along its route.

Isoclinics and isodynamics intersect at smaller angles in the North Atlantic. Nevertheless, the two sets of isolines form a non-orthogonal grid that could be used to approximate global position from nearly all locations along the North Atlantic gyre (Fig. 4), as well as from all known loggerhead feeding grounds along the North American coast²².

If a magnetic map is indeed used by sea turtles during navigation, it might range from a system that provides only a crude estimate of global position to one capable of pinpointing a specific beach. The resolution would depend in part on the sensitivity of the turtles to the two magnetic parameters, which at present is entirely unknown. However, honeybees may be able to detect intensity differences of as little as 26 nT (ref. 23), and a sensitivity of approximately 10–20 nT has been hypothesized for birds². If turtles have similar abilities to discriminate intensity, and can also detect inclination angle with great precision, a highly accurate system is at least theoretically possible. Alternatively, however, a magnetic map might function only to guide turtles into the general vicinity of a nesting beach, close enough for other cues (for example, olfactory or visual) to take over and allow final localization of the nesting site.

The bicoordinate magnetic maps we have proposed could presumably not provide accurate positional fixes at sites of strong magnetic anomalies, and the exact east–west position cannot be unambiguously determined in those few parts of the central North Atlantic where isoclinics and isodynamics are parallel (Fig. 4). Because areas with coinciding isolines are relatively restricted, however, and because none occur in locations frequented by adult turtles, we do not consider them to pose a serious problem for the hypothesis. Moreover, migrating through such regions would not necessarily be difficult; a turtle on its way to a North American nesting beach might, for example, simply swim westwards until it re-enters an area where an accurate fix of position is again possible.

Even if turtles do not possess or use bicoordinate magnetic maps, the ability to detect subtle differences in field intensity may still be used in long-distance navigation. For example, turtles might detect magnetic anomalies and use them as landmarks for piloting through familiar areas, or they might follow magnetic maxima and minima pathways that have arisen as a result of seafloor spreading. Some whales and dolphins are thought to swim along such magnetic lineations while migrating^{24–26}.

Thus the ability to detect magnetic field intensity might be used by sea turtles in several different ways during long-distance

migrations. For now, however, precisely how sea turtles navigate in the open ocean remains an enduring mystery of behavioural biology. □

Received 8 November 1995; accepted 10 January 1996.

1. Viguier, C. *Rev. Philos. France Etranger* **14**, 1–36 (1882).
2. Gould, J. L. in *Magnetite Biomineralization and Magnetoreception in Organisms* (eds Kirschvink, J. L., Jones, D. A. & McFadden, B. J.) 257–268 (Plenum, New York, 1985).
3. Baker, R. R. *Bird Navigation: the Solution of a Mystery?* (Holmes and Meier, New York, 1984).
4. Walcott, C. in *Orientation in Birds* (ed. Berthold, P.) 38–51 (Birkhäuser, Boston, MA, 1991).
5. Berthold, P. *Bird Migration* (Oxford Univ. Press, 1993).
6. Wallraff, H. G. *Comp. Biochem. Physiol.* **76A**, 643–663 (1983).
7. Meylan, A. B., Bowen, B. W. & Avise, J. C. *Science* **248**, 724–727 (1990).
8. Bowen, B. W. *et al. Conserv. Biol.* **7**, 834–844 (1993).
9. Bowen, B. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **92**, 3731–3734 (1995).
10. Lohmann, K. J. & Lohmann, C. M. F. *J. exp. Biol.* **194**, 23–32 (1994).
11. Carr, A. *The Sea Turtle: So Excellent a Fish* (University of Texas Press, Austin, 1986).
12. Carr, A. *Copeia* **3**, 547–555 (1975).
13. Carr, A. in *Bio-Telemetry* (ed. Slater, L. E.) 179–193 (MacMillan, New York, 1963).
14. Limpus, C. J. *et al. Wildl. Res.* **19**, 347–358 (1992).
15. Papi, R., Liew, H. C., Luschi, P. & Chan, E. H. *Mar. Biol.* **122**, 171–175 (1995).
16. Salmon, M. & Wyneken, J. *Herp. nat. Hist.* **2**, 13–24 (1994).
17. Carr, A. *Conserv. Bio.* **1**, 103–121 (1987).
18. Carr, A. *Bioscience* **36**, 92–100 (1986).
19. Lohmann, K. J. *J. exp. Biol.* **155**, 37–49 (1991).
20. Lohmann, K. J. & Lohmann, C. M. F. *J. exp. Biol.* **190**, 1–8 (1994).
21. Batschelet, E. *Circular Statistics in Biology* (Academic, London, 1981).
22. Dodd, C. K. Jr *U.S. Fish Wildl. Serv. Biol. Rep.* **88**, 1–110 (1988).
23. Walker, M. M. & Bitterman, M. E. *J. exp. Biol.* **145**, 489–494 (1989).
24. Klinowska, M. *Aquat. Mamm.* **1**, 27–32 (1985).
25. Kirschvink, J. L., Dizon, A. E. & Westphal, J. A. *J. exp. Biol.* **120**, 1–24 (1986).
26. Walker, M. M., Kirschvink, J. L., Ahmed, G. & Dizon, A. E. *J. exp. Biol.* **171**, 67–78 (1992).
27. U.S. Defense Mapping Agency Hydrographic/Topographic Center *Magnetic Inclination of Dip, Epoch 1985.0* (U.S. Government Printing Office, Washington DC, 1985).
28. U.S. Defense Mapping Agency Hydrographic/Topographic Center *Total Intensity of the Earth's Magnetic Field, Epoch 1985.0* (U.S. Government Printing Office, Washington DC, 1985).

ACKNOWLEDGEMENTS. We thank K. Jackson, J. Wang, M. Salmon and J. Wyneken for assistance. This work was supported by grants from the National Science Foundation.

Role of protozoan grazing in relieving iron limitation of phytoplankton

Katherine Barbeau, James W. Moffett*, David A. Caron, Peter L. Croot & Deana L. Erdner

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

RECENT evidence indicates that iron is a limiting factor in primary production in some areas of the oceans^{1,2}. In sea water, iron is largely present in the form of particulate and colloidal phases which are apparently unavailable for uptake by phytoplankton^{3–5}. Several mechanisms have been proposed whereby non-reactive iron may be converted into more labile forms (for example, thermal dissolution⁶, photochemical reactions^{7,8} and ligand complexation⁹). Here we report that digestion of colloidal iron in the acidic food vacuoles of protozoan grazers may be a mechanism for the generation of 'bioavailable' iron from refractory iron phases. We have demonstrated several grazer-mediated effects on colloidal ferrihydrite, including a decrease in colloid size, an increase in colloid lability as determined by competitive ligand-exchange techniques, and an increase in the bioavailability of colloids to iron-limited diatoms. These results indicate that protozoan grazers may significantly enhance the supply of iron to marine phytoplankton from terrestrial sources.

Heterotrophic nanoprotozoans are known to be important components of the marine planktonic ecosystem as the primary consumers of bacterial biomass¹⁰ and recyclers of major nutrients^{11,12}. These grazers consume particles and colloids in

* To whom correspondence should be addressed.

the 0.2–1- μm size class¹³, which in sea water contains a significant pool of particulate and colloidal iron⁹. Once incorporated into the grazer food vacuole, these iron species may experience acidic conditions (pH 2–3) and intense enzymatic activity¹⁴ during the digestive process. Dissolution of iron oxyhydroxides and dissociation of organically complexed iron is likely to occur, resulting in the formation of more labile, bioavailable forms such as dissolved Fe(III) hydroxy species.

To test this hypothesis, we have conducted a series of experiments in simplified model systems containing protozoan grazers, bacteria, phytoplankton and iron colloids. The colloids used in our experiments were a modified version of the 50° ferrihydrite described in ref. 3. Electron micrographs of the concentrated colloidal sol revealed a structure of filamentous aggregates (~150 nm in length) of somewhat irregularly shaped and sized subunits (~10 nm in size). The colloids were added to sea water at iron concentrations of approximately 1 μM .

Figure 1 shows the results of a size-fractionation experiment using colloids that were radiolabelled by adding $^{59}\text{Fe}(\text{III})$ during ferrihydrite synthesis. The experimental system consisted of two replicates each of a colloid control (a seawater suspension of colloids), a bacteria control (bacteria and colloid suspension) and the experimental culture (bacteria and colloid suspension inoculated with protists). The protist used in this experiment, *Cafeteria* sp., is a common planktonic heterotrophic bicosoecid microflagellate, about 2–4 μm in size¹⁵. In the presence of the grazer,

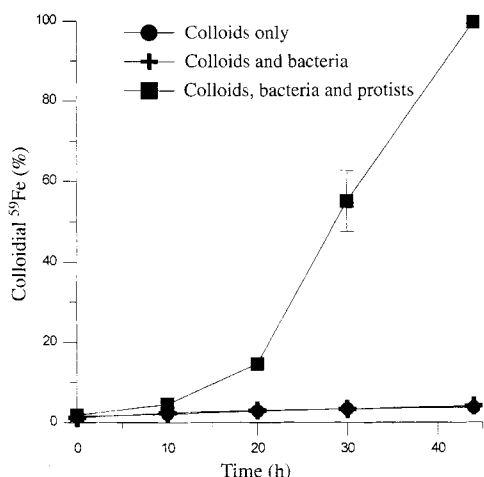


FIG. 1 Results of experiment with ^{59}Fe -labelled colloids, showing the percentage of the colloidal iron that passes through a 50-nm filter over time. Ferrihydrite was synthesized according to ref. 3, and the resulting colloidal sol was briefly extracted (15 min) in a slurry of Biorad 100 mesh cation exchange resin at ambient pH (~3) to remove dissolved or highly labile iron. 0.2 μm filtered, autoclaved coastal seawater (background iron concentration 21 nM) was used in these experiments. 10^{-4} M EDTA was added, but subsequent experiments indicated that the presence of EDTA did not affect experimental results. The bacterium, *Halomonas halodurans*, was grown separately in a yeast extract culture and then centrifuged and rinsed several times in sterile sea water before being resuspended (concentration 8×10^7 cells ml^{-1}) with colloidal ferrihydrite in bacteria control and experimental bottles. An inoculum of *Cafeteria* was added to experimental bottles at the start of the experiment (concentration 7.3×10^3 cells ml^{-1}). Model systems were maintained in polycarbonate bottles at room temperature under normal laboratory lighting, with no agitation. Size fractionation was done by removing aliquots of culture at several time-points during the 45-h incubation and filtering them through 0.05 μm Nuclepore polycarbonate filters using a syringe filtration set-up. This method has been shown to minimize damage to fragile protozoan membranes²². Filtrates were counted for ^{59}Fe activity on a Canberra low-energy germanium detector. All filtrate counts were normalized to total ^{59}Fe counts on an aliquot of culture taken at the same time as the filtered aliquots to correct for wall losses, which were <5%. Error bars are s.d. of two replicate cultures.

100% of the ferrihydrite was converted into smaller colloids and/or dissolved species less than 50 nm in size within 45 hours. By this time the protists had consumed a large portion of the bacteria, as indicated by the decreased optical density of the experimental culture relative to the bacteria control.

In conjunction with size fractionation, changes in colloid lability

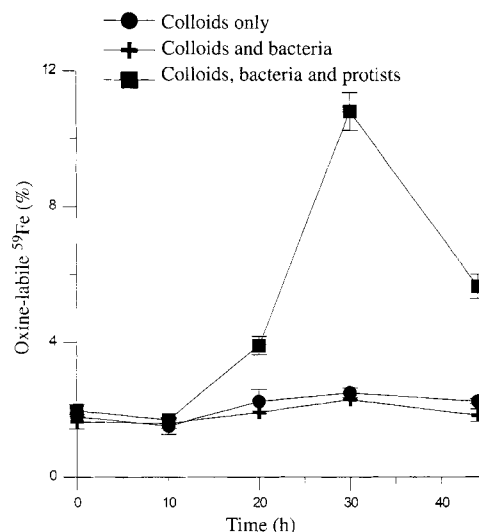


FIG. 2 Oxine lability measurements for the same model systems discussed in Fig. 1. Oxine lability was determined according to ref. 3. Whole-culture aliquots were incubated with oxine for 1 h, then prefiltered through a 0.2 μm Nuclepore (syringe with Swinnex holder, see Fig. 1) and the filtrate passed through a C-18 Sep-Pak cartridge. The methanol eluate from each cartridge was counted for ^{59}Fe activity, and the percentage of total ^{59}Fe determined by normalizing to the activity of a whole culture sample taken at the same time. Error bars are the s.d. of two replicate cultures.

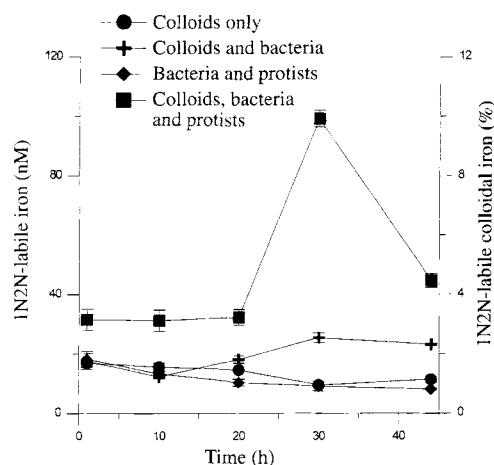


FIG. 3 Results of chemical lability assays using the CSV with 1N2N method as in ref. 17. Whole-culture aliquots were incubated with the added ligand before CSV analysis. Model systems were different from those described in Figs 1 and 2, but were identical except for the absence of any added EDTA and the use of bacteria (*Halomonas halodurans*) grown on a glucose-based medium. Concentrations of bacteria and protists at t_0 were similar to those given for Figs 1 and 2. As this method does not involve radiolabelled iron colloids, an additional control containing bacteria, *Cafeteria*, and no colloids was necessary to show that the increase in labile iron is due to iron from the ferrihydrite colloids, not from the bacteria. The left y-axis expresses the concentration of 1N2N-labile iron in the samples, determined by standard additions. In the right y-axis, the concentration of 1N2N-labile iron has been normalized to the concentration of added colloidal iron (~1 μM). The right axis is directly comparable to the y-axis in Fig. 2. Error bars are the s.d. of replicate cultures.

in the same model systems were assessed by determining the reactivity of the ferrihydrite with a synthetic iron chelator, oxine. A correlation between oxine exchangeable iron and biologically available iron has been demonstrated^{3,16}. The results of oxine lability determinations are shown in Fig. 2. The culture containing *Cafeteria* exhibits approximately a fivefold increase in the percent-

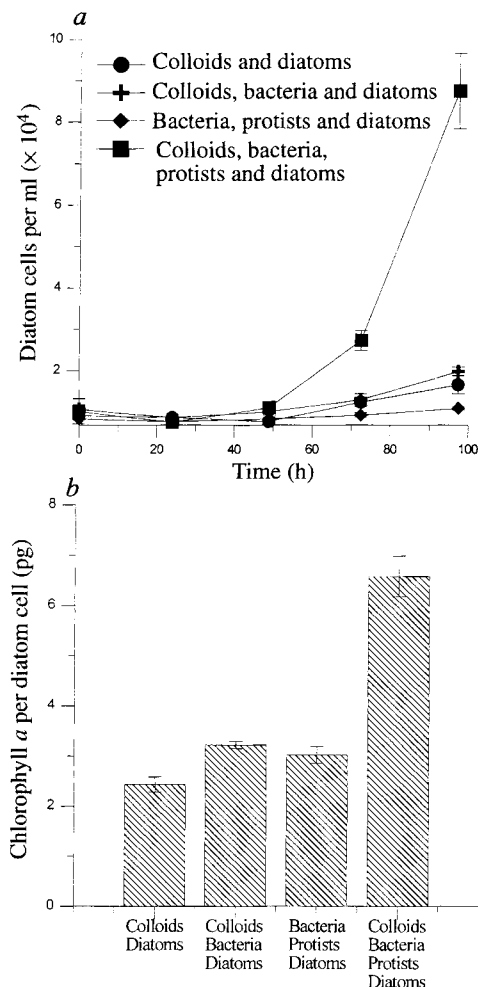


FIG. 4 Results of experiments with iron-limited diatoms. *a*, Diatom cell numbers and *b*, chlorophyll concentrations per cell. Cultures were set up in triplicate in filtered, sterilized coastal seawater with f/2 nutrients (except iron) added (background iron concentration 21 nM). 1 μ M EDTA was also added. The bacteria were grown in a yeast extract medium, heat-killed (2 h at 60 °C), rinsed, and resuspended in the model systems. Colloid controls contained diatoms grown in iron-limiting media and resuspended with colloidal ferrihydrite. Bacteria controls contained diatoms, colloids, and bacteria in suspension. An additional control (diatoms and bacteria inoculated with protists) was necessary to detect enhancement in diatom growth/chlorophyll due to remineralization of bacterial biomass by protists. Experimental cultures contained diatoms, bacteria, and iron colloids inoculated with protists. Bacteria and protist cell concentrations at t_0 were similar to those given for Figs 1 and 2. The protistan inoculum had been transferred into heat-killed bacteria over 10 times. Cultures were maintained without shaking in an incubator at 20 °C, illuminated by a cool-white fluorescent light at $\sim 100 \mu\text{Einsteins m}^{-2} \text{s}^{-1}$ on a 16 h light:8 h dark schedule. Triplicate samples for diatom-cell counts were taken (after thorough mixing) roughly every 24 h from t_0 . Diatom-cell concentrations were determined by light microscopy using a hemacytometer. Chlorophyll *a* concentrations were determined by standard fluorometric methods at 96 h on triplicate samples. Error bars are the s.d. of triplicate cultures. Samples preserved in 1% glutaraldehyde were examined for the presence of chlorophyll fluorescence in *Cafeteria* food vacuoles to make certain that the grazer (2–4 μm in size) was not consuming any diatoms (>10 μm in size). No fluorescing *Cafeteria* were found.

tage of oxine-labile iron relative to the control cultures, peaking at 30 hours.

The chelator 1-nitroso-2-naphthol (1N2N) was used to study changes in ferrihydrite reactivity (Fig. 3). Cathodic stripping voltammetry (CSV) was used to determine the concentration of 1N2N-labile iron^{17,18}. As with the oxine technique, the experimental culture demonstrates a fivefold increase in labile iron, peaking at 30 hours. This coincides with peak rates of protozoan grazing, when decreases in optical density of the grazer-containing cultures first become apparent. We suggest that as grazing rates subsequently decline, owing to reduced prey density, the pool of labile iron is decreased by rapid reactions to form more refractory phases. The results of Fig. 1 do not demonstrate the same time dependency, presumably because the iron does not re-aggregate into size classes larger than 50 nm within the time scale of our experiments.

To assess directly the impact of protozoan grazers on the bioavailability of colloidal iron, we did an experiment using model systems containing heat-killed bacteria (*Halomonas halodurans*), protists (*Cafeteria* sp.), ferrihydrite, and iron-limited diatoms (*Thalassiosira weissflogii*, clone Actin) (Fig. 4). In the cultures containing diatoms in association with both ferrihydrite and protists, final diatom cell densities were eight times higher and concentrations of chlorophyll *a* per diatom cell were twice as high as in control cultures. This indicates production of bioavailable iron from non-labile ferrihydrite through the grazing activities of protists. The time dependence of the effect is determined by the physiological response time of the diatoms. The bacteria in this experiment were heat-killed in order to demonstrate that grazer-mediated release of bioavailable iron did not result from previous uptake of colloidal iron by the bacteria.

Additional data (not shown) demonstrate that the effect of the grazers on iron colloid lability is independent of medium pH. We have also found that the medium of an actively grazing protozoan culture from which protists have been removed by 1- μm filtration does not affect ferrihydrite lability (data not shown). Our data support the conclusion that the effect of protozoan grazers on iron colloids is due to the ingestion and chemical transformation of the colloids within the low-pH, enzyme-rich microenvironment of the protozoan food vacuole.

Extrapolation from our model systems to the diverse assemblages of particles and microorganisms in the oceans is premature. However, a rough approximation made by multiplying average clearance rates ($2 \times 10^{-6} \text{ ml cell}^{-1} \text{ h}^{-1}$)¹⁹ of microflagellates by typical concentrations ($10^3 \text{ cells per ml}$)¹⁹ of microflagellates in open ocean waters indicates that protists may potentially produce bioavailable iron from colloidal Fe(III) in the oceans with a first-order rate constant of approximately $2 \times 10^{-3} \text{ h}^{-1}$. This is a tenth of the maximum (noon-time) near-surface first-order rate constant for photoreduction of colloidal iron, $1.8 \times 10^{-2} \text{ h}^{-1}$ (refs 8, 20). But photoreduction rate declines exponentially with depth as a result of light attenuation²¹, and is also well below the maximum for much of the day and on cloudy days, as well as at night. Therefore, when integrated temporally over the entire euphotic zone, protozoan grazing may equal or exceed photoreduction in increasing iron availability to phytoplankton. □

Received 25 October 1995; accepted 8 January 1996.

- Martin, J. H. & Fitzwater, S. E. *Nature* **331**, 341–343 (1988).
- Martin, J. H. et al. *Nature* **371**, 123–129 (1994).
- Wells, M. L., Mayer, L. M. & Guillard, R. R. L. *Mar. Chem.* **33**, 23–40 (1991).
- Rich, H. W. & Morel, F. M. M. *Limnol. Oceanogr.* **35**, 652–662 (1990).
- Martin, J. H., Gordon, R. M., Fitzwater, S. & Broenkow, W. W. *Deep Sea Res.* **36**, 649–680 (1989).
- Wells, M. L., Zorkin, N. S. & Lewis, A. G. *J. mar. Res.* **41**, 731–746 (1983).
- Waite, D. T. & Morel, F. M. M. *Environ. Sci. Technol.* **18**, 860–868 (1984).
- Wells, M. L. & Mayer, L. M. *Deep Sea Res.* **38**, 1379–1395 (1991).
- Wells, M. L., Price, N. M. & Bruland, K. W. *Mar. Chem.* **48**, 157–182 (1995).
- Azam, F. T. et al. *Mar. Ecol. Prog. Ser.* **10**, 257–263 (1983).
- Caron, D. A., Goldman, J. C. & Dennett, M. R. *Hydrobiologia* **159**, 27–40 (1988).
- Caron, D. A. in *Protozoa and their Role in Marine Processes* NATO ASI Series Vol. G 25 (eds Reid, P. C., Turley, C. M. & Burkhill, P. H.) 387–416 (Springer, Berlin, 1990).
- Fenchel, T. *Ecology of Protozoa: The Biology of Free-living Phagotrophic Protists* (Science Tech, Madison, 1987).

14. Fok, A. K., Lee, Y. & Allen, R. D. *Protozool.* **29**, 409–414 (1982).
15. Fenchel, T. & Patterson, D. J. *Mar. Micr. Food Webs* **3**, 9–19 (1988).
16. Wells, M. L., Mayer, L. M. & Guillard, R. R. L. *Mar. Ecol. Prog. Ser.* **69**, 93–102 (1991).
17. Wu, J. & Luther, G. W. *Mar. Chem.* **50**, 159–178 (1995).
18. Wu, J. & Luther, G. W. *Limnol. Oceanogr.* **39**, 1119–1129 (1994).
19. Beminger, U.-G., Caron, D. A., Sanders, R. W. & Finlay, B. J. in *The Biology of Free-living Heterotrophic Flagellates*. Systematics Association Special Vol. 45 (eds Patterson, D. J. & Larsen, J.) 39–56 (Clarendon, Oxford, 1991).
20. Voelker, B. M. & Sedlak, D. L. *Mar. Chem.* **50**, 93–102 (1995).
21. Wells, M. L. et al. *Nature* **353**, 248–250 (1991).
22. Nagata, T. & Kirchman, D. L. *Mar. Ecol. Prog. Ser.* **68**, 1–5 (1990).

ACKNOWLEDGEMENTS. We thank D. Anderson for use of laboratory facilities, D. Kulis for technical assistance, and E. Sholkovitz and B. Voelker for helpful comments on this manuscript. Financial support was provided by the US Office of Naval Research and the National Science Foundation.

Sheep cloned by nuclear transfer from a cultured cell line

K. H. S. Campbell, J. McWhir, W. A. Ritchie & I. Wilmut

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK

NUCLEAR transfer has been used in mammals as both a valuable tool in embryological studies¹ and as a method for the multiplication of 'elite' embryos^{2–4}. Offspring have only been reported when early embryos, or embryo-derived cells during primary culture, were used as nuclear donors^{5,6}. Here we provide the first report, to our knowledge, of live mammalian offspring following nuclear transfer from an established cell line. Lambs were born after cells derived from sheep embryos, which had been cultured for 6 to 13 passages, were induced to quiesce by serum starvation before transfer of their nuclei into enucleated oocytes. Induction of quiescence in the donor cells may modify the donor chromatin structure to help nuclear reprogramming and allow development. This approach will provide the same powerful opportunities for analysis and modification of gene function in livestock species that are available in the mouse through the use of embryonic stem cells⁷.

The cells used in these experiments were isolated by microdissection and explantation of the embryonic disc (ED) of day 9 *in vivo* produced 'Welsh mountain' sheep embryos. The line was established from early passage colonies with a morphology like that of embryonic stem (ES) cells. By the second and third passages, the cells had assumed a more epithelial, flattened morphology (Fig. 1a) which was maintained on further culture (to at least passage 25). At passage 6, unlike murine ES cells they expressed cytokeratin, and nuclear lamin A/C which are markers associated with differentiation⁸. This embryo-derived epithelial cell line has been designated TNT4 (for totipotent for nuclear transfer).

The development of embryos reconstructed by nuclear transfer is dependent upon interactions between the donor nucleus and the recipient cytoplasm. We have previously reported the effects of the cytoplasmic kinase activity, maturation/mitosis/meiosis promoting factor (MPF), on the incidence of chromosomal damage and aneuploidy in reconstructed embryos and established two means of preventing such damage⁹. First, the effects of the donor cell-cycle stage can be overcome by transferring nuclei after the disappearance of MPF activity by prior activation of the recipient enucleated MII oocyte^{9,10}. Using this approach we obtained the birth of lambs by nuclear transfer during establishment of the cell line (up to and including passage 3). On subsequent culture (passages 6 and 11) no development to term was obtained (see Table 1). From these numbers we cannot conclude that development to term will not be obtained using

this method. The lack of development of some control embryos is thought to relate to an infection in the oviduct of the temporary recipient ewe from which 6 were recovered.

An alternative means of avoiding damage due to the activity of MPF is to transfer diploid nuclei into metaphase II oocytes that have a high level of MPF activity⁹. The availability of TNT4 cells allows this approach to be used. In this study a synchronous population of diploid donor nuclei was produced by inducing the cells to exit the growth cycle and arrest in G0 in a state of quiescence. In the presence of a high level of MPF activity the transferred nucleus undergoes nuclear membrane breakdown and chromosome condensation. It has been argued¹¹ that the developmental potential of reconstructed embryos depends upon the "reprogramming of gene expression" by the action of cytoplasmic factors and that this might be enhanced by the prolongation of this period of exposure. To assess these effects donor cells were fused to oocytes either (1) 4–8 h before activation 'post-activated' or (2) at the time of activation 'fusion and activation' or (3) to pre-activated oocytes 'preactivated'.

During these studies *in vivo* ovulated metaphase arrested (MII) oocytes were flushed from the oviduct of 'Scottish blackface' ewes. The methodology used was as previously described¹⁰ with the following exceptions; oocytes were recovered 28–33 h after injection of gonadotropin-releasing hormone (GnRH), calcium/magnesium-free PBS containing 1.0% FCS was used for all flushing, and recovered oocytes were transferred to calcium-free M2 medium¹² containing 10% FCS and were maintained at 37 °C in 5% CO₂ in air until use. As soon as possible after recovery oocytes were enucleated and embryos reconstructed. At 50–54 h

TABLE 1 Development using unsynchronized TNT4 cells

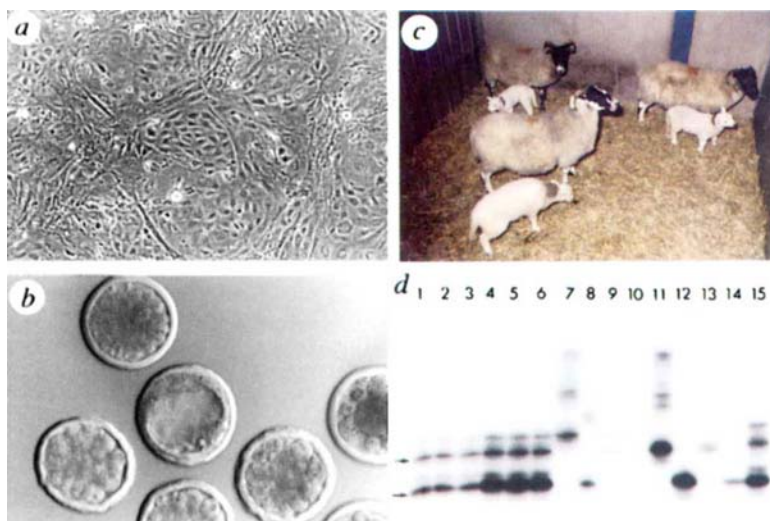
Donor cell type	Number of morula and blastocysts/total embryos (%)	Number of lambs/embryos transferred
October 1993–February 1994		
16 cell	6/11 (27.3)	2/6
ED cell	1/15 (6.7)	0/1
ED P1	4/19 (21.0)	1/4
ED P2	1/11 (9.1)	1/1
ED P3	2/36 (5.5)	2/2
October–December 1994		
16 cell	14/28 (50.0)	0/14
TNT4 P6	9/98 (9.2)	0*/9
TNT4 P11	10/92 (10.9)	0/10

Development of ovine embryos reconstructed by nuclear transfer of unsynchronized cells during isolation and after establishment of the TNT4 line to enucleated preactivated ovine oocytes. P, Passage number; ED, embryonic disc. For embryo reconstruction, donor oocytes were placed into calcium-free M2 containing 10% FCS, 7.5 µg ml⁻¹ Cytochalasin B (Sigma) and 5.0 µg ml⁻¹ Hoechst 33342 (Sigma) at 37 °C for 20 min to aspirate. A small amount of cytoplasm enclosed in plasma membrane was removed from directly beneath the 1st polar body using a glass pipette (~20 µm tip external diameter). Enucleation was confirmed by exposing this karyoplast to ultraviolet light and checking for the presence of a metaphase plate. At 34–36 h after GnRH injection enucleated oocytes were activated. Following further culture for 4–6 h in TC199, 10% FCS a single cell was fused. All activations and fusions were accomplished as previously described^{10,17} in 0.3 M mannitol, 0.1 mM MgSO₄, 0.0005 mM CaCl₂¹⁷. For activation a single DC pulse of 1.25 kV cm⁻¹ for 80 µs and for fusion an AC pulse of 3 V for 5 s followed by 3 d.c. pulses of 1.25 kV cm⁻¹ for 80 µs were applied. All oocyte/cell couplets were cultured in TC199, 10% FCS 7.5 µg ml⁻¹ Cytochalasin B (SIGMA) for 1 h following application of the fusion pulse and then in the same medium without Cytochalasin until transferred to temporary recipient ewes. Reconstructed embryos were cultured in the ligated oviduct of a recipient 'blackface' ewe until day 7 after reconstruction. All morula and blastocyst stage embryos were transferred to synchronized recipient blackface ewes for development to term.

* A single pregnancy was established but subsequently lost at about 70–80 days.

FIG. 1 Production and characterization of the TNT4 cell line and the offspring produced by nuclear transfer from TNT4 cells. *a*, Morphology of the TNT4 cell line at passage 6. *b*, Group of embryos including a single blastocyst on day 7 after reconstruction. *c*, Group of three Welsh mountain lambs produced by nuclear transfer with surrogate Scottish black-face ewes. *d*, Autoradiogram showing the alleles generated following amplification of the microsatellite FCB266 (ref. 18). Lanes 1–6 are from, respectively, TNT4 cells and the five lambs generated by nuclear transfer. Both lambs and cells display an identical pattern, revealing 2 alleles (arrowed) at 114 and 125 bp. Lanes 7–15, nine randomly chosen Welsh mountain sheep, none of whom show an identical pattern to the nuclear transfer group. Lambs and TNT4 cells were also identical at six further microsatellite loci: MAF33, MAF48, MAF65, MAF209, OarFCB11, OarFCB128, OarRCB304 (data not shown). The nine unrelated random control animals showed extensive variation at all of these loci.

METHODS. Groups of 4–6 microdissected embryonic discs were cultured on feeder layers of mitotically inactivated primary murine fibroblasts in Dulbeccos Modified Eagles medium (GIBCO) containing 10% fetal calf serum, 10% newborn serum and supplemented with recombinant human leukaemia inhibition factor (LIF). After 5–7 days of culture, expanding discs were treated with trypsin and passaged onto fresh feeders yielding 4 similar lines. At passage 12 of the 2n chromosome complement of 54 was observed in 31 of 50 spreads, the remaining aneuploid spreads are thought to be artefacts of preparation. For microsatellite analysis genomic



DNA was extracted from whole blood, tissue culture cells or fetal tissues using a puregene DNA isolation kit (Gentra Systems Inc., Minneapolis, USA). The PCR analysis of microsatellites was carried out using an end-labelled primer ($[\gamma\text{-}^{32}\text{P}]\text{ATP}$). All other aspects of labelling and thermal cycling conditions were as described elsewhere¹⁷.

after GnRH injection, reconstructed embryos were embedded in agar and transferred to the ligated oviduct of dioestrus ewes. After 6 days the embryos were retrieved and development assessed microscopically (see Fig. 1*b*).

The development of embryos reconstructed using quiescent TNT4 cells and 3 different cytoplasm recipients is summarized in Table 2. No significant difference was observed in the frequency of development with high and low passage number donor cells or with cytoplasm recipient type used (results were analysed by the marginal model in ref. 13). All embryos that had developed to the morula/blastocyst stage were transferred as soon as possible to the uterine horn of synchronized final recipient ewes for development to term. Recipient ewes were monitored for pregnancy by ultra-

sonography. Ewes that were positive at day 35 were classified as pregnant (Table 3). A total of eight fetuses were detected in seven recipient ewes including a single twin pregnancy. A total of five phenotypically female Welsh mountain lambs were born from the Scottish blackface recipient ewes (Fig. 1*c*). Two of these lambs died within minutes of birth and a third at 10 days; the remaining two lambs are apparently normal and healthy (8–9 months old). Of the remaining 3 fetuses, one was lost at about 80 days of gestation, and a second was lost at 144 days of gestation. The third fetus was thought to be a twin pregnancy and was either misdiagnosed or lost at an unknown time. Microsatellite analysis of the cell line, fetuses and lambs showed that all of the female lambs were derived from a single cell population (Fig. 1*d*).

TABLE 2 Development to morula and blastocyst stage of ovine embryos reconstructed using quiescent TNT4 cells and 3 different cytoplasm recipients (January–March 1995)

Experiment number	Cytoplasm type TNT passage number	Number of morulae and blastocysts/total number of embryos recovered (%)		
		Post-activated	Activation and fusion	Preactivated
1	6	4/28	6/32*	–
2	7	1/10	1/26*	–
3	13	0/2	–	2/14
4	13	0/14	0/11	–
5	11	1/9	–	0/9
6	11	1/2	9/29***	–
7	12	–	–	6/45*
8	13	3/13*	–	–
Total		10/78 (12.8%)	16/98 (16.3%)	8/68 (11.7%)

Development to the morula and blastocyst stage of ovine embryos recovered on day 7 after reconstruction by nuclear transfer of quiescent TNT4 cells at different passages into 3 cytoplasm recipients. To induce quiescence, TNT4 cells were plated into feeder layers in 29-cm² flasks (GIBCO) and cultured for 2 days, the semiconfluent exponentially growing cultures were then washed three times in medium containing 0.5% FCS and cultured in this low-serum medium for 5 days. Embryos were reconstructed using preactivated cytoplasts as previously described (Table 1) and by two other protocols. (1) post-activation, as soon as possible after enucleation a single cell was fused to the cytoplasm in 0.3 M mannitol without calcium and magnesium, to prevent activation. Couplets were washed and cultured in calcium-free M2, 10% FCS at 37 °C, 5% CO₂ for 4–8 h. Thirty minutes before activation the couplets were transferred to M2 medium, 10% FCS containing 5 μM Nocodazole (SIGMA). Following activation the reconstructed zygotes were incubated in medium TC199, 10% FCS, 5.0 μM Nocodazole for a further 3 h. (2) Preactivation, at 34–36 h after GnRH injection a single cell was fused to an enucleated oocyte. The same pulse also induced activation of the recipient cytoplasm. All activations and fusions were accomplished as described in Table 1 unless otherwise stated.

* Denotes number of pregnancies following transfer of morula and blastocyst stage embryos to synchronized final recipient ewes.

TABLE 3 Induction of pregnancy and further development following transfer of morula and blastocyst stage embryos reconstructed from quiescent TNT4 cells

Cytoplasm type	Post-activated	Activation and fusion	Preactivated
Total number of morula and blastocyst stage embryos transferred	10	16	8
Total number of ewes	6	9	4
Number of pregnant ewes (%)	1 (16.7)	5 (55.5)	1 (25.0)
Number of fetuses/total embryos transferred (%)	2/10 (20.0)	5/16 (31.25)	1/8 (12.5)
Number of live births	1	3	1
Passage number of cells resulting in offspring	1 × P11	1 × P6, 2 × P11	1 × P13

Induction of pregnancy following transfer of all morula/blastocyst stage reconstructed embryos to the uterine horn of synchronized final recipient blackface ewes. The table shows the total number of embryos from each group transferred, the frequency of pregnancy in terms of ewes and embryos (in the majority of cases 2 embryos were transferred to each ewe and a single twin pregnancy was established (using the 'post-activated' cytoplasm) and the number of live lambs obtained.

Because of the seasonality of sheep a direct comparison of all of these methods of embryo reconstruction has not yet been made. The success of the later studies may be due to a number of factors. First, quiescent nuclei are diploid and therefore the cell-cycle stages of the karyoplast and cytoplasm in both the 'post-activation' and 'fusion and activation' methods of reconstruction are coordinated. The preactivated cytoplasm will accept donor nuclei from G0, G1, S and G2 cell-cycle phases. Second, the G0 phase of the cell cycle has been implicated in the differentiation process and the chromatin of quiescent nuclei has been reported to undergo modification¹⁴. As a result the chromatin of quiescent donor nuclei may be more readily modified by oocyte cytoplasm. The TNT4 cells resemble several cell lines derived previously in sheep¹⁵ and also pigs¹⁶. It remains to be determined whether comparable development is obtained with other such lines or other cell types. At the present time we are unable to differentiate the mechanisms involved and report that the combination of nuclear transfer and cell type described here support development to term of cloned ovine embryos from cells that had been in culture through up to 13 passages. As cell-cycle duration was about 24 h, this period of culture before nuclear transfer would be sufficient to allow genetic modification and selection if procedures comparable to those used in murine ES cells can be established.

The production of cloned offspring in farm animal species could provide enormous benefits in research, agriculture and biotechnology. The modification by gene targeting and selection of cell populations before embryo reconstruction coupled to the clonal origin of the whole animal provides a method for the dissemination of rapid genetic improvement and/or modification into the population □

Received 27 November 1995; accepted 5 January 1996.

- McGrath, J. & Solter, D. *Science* **220**, 1300–1302 (1983).
- Bondiolli, K. R., Westhusin, M. E. & Looney, C. R. *Therio* **33**, 165–174 (1990).
- Prather, R. S. & First, N. L. *Int. Rev. Cytol.* **120**, 169–190 (1990).
- Chesne, P., Heyman, Y., Peynot, N. & Renard, J.-P. *C.R. Acad. Sci. Paris Life Sci.* **316**, 487–491 (1993).
- Sims, M. & First, N. L. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6243–6147 (1994).
- Collas, P. & Barnes, F. L. *Molec. Reprod. Dev.* **38**, 264–267 (1994).
- Hooper, M. L. *Embryonal Stem Cells: Introducing Planned Changes into the Germline* (ed. Evans, H. J.) (Harwood Academic, Switzerland, 1992).
- Galli, C., Lazzari, G., Flechon, J. & Moor, R. M. *Zygote* **2**, 385–389 (1994).
- Campbell, K. H. S., Ritchie, W. A. & Wilmut, I. *Biol. Reprod.* **49**, 933–942 (1993).
- Campbell, K. H. S., Loi, P., Capai, P. & Wilmut, I. *Biol. Reprod.* **50**, 1385–1393 (1994).
- Szollasi, D., Czolowska, R., Szollasi, M. S. & Tarkowski, A. K. *J. Cell Sci.* **91**, 603–613 (1988).
- Whitten, W. K. & Biggers, J. D. *J. Reprod. Fertil.* **17**, 399–401 (1968).
- Breslaw, N. E. & Clayton, D. G. *J. Am. Stat. Assoc.* **88**, 9–25 (1993).
- Whitfield, J. F., Boynton, A. L., Rixon, R. H. & Youdale, T. *Control of Animal Cell Proliferation Vol. 1* (eds Boynton, A. L. & Lefert, H. L.) 331–365 (Academic, London 1985).
- Piedrahita, J. A., Anderson, G. B. & Bon Durrant, R. H. *Therio* **34**, 879–901 (1990).
- Gerfen, R. W. & Wheeler, M. B. *Anim. Biotechnol.* **6**, 1–14 (1995).
- Willadsen, S. M. *Nature* **320**, 63–65 (1986).
- Buchanan, F. C., Galloway, S. M. & Crawford, A. M. *Anim. Genet.* **25**, 60 (1994).

ACKNOWLEDGEMENTS. We thank M. Ritchie, J. Bracken, M. Malcolm-Smith and R. Ansell for technical assistance; D. Waddington for statistical analysis; and H. Bowran and his colleagues for their care of the animals. This research was supported by Ministry of Agriculture Fisheries and Food and the Roslin Institute.

Midbrain development induced by FGF8 in the chick embryo

Philip H. Crossley, Salvador Martinez* & Gail R. Martin†

Department of Anatomy and Program in Developmental Biology, School of Medicine, University of California at San Francisco, San Francisco, California 94143-0452, USA

* Department of Morphological Sciences, Faculty of Medicine, University of Murcia, Institute of Neurosciences, University of Alicante, Spain

VERTEBRATE midbrain development depends on an organizing centre located at the isthmus, a constriction in the embryonic mid/hindbrain region^{1–3,28}. Isthmic tissue grafts transform chick caudal forebrain into an ectopic midbrain that is the mirror image of the normal midbrain⁴. Here we report that FGF8 protein has the same midbrain-inducing and polarizing effect as isthmic tissue. Moreover, FGF8 induces ectopic expression in the forebrain of genes normally expressed in the isthmus, suggesting that the ectopic midbrain forms under the influence of signals from a new 'isthmus-like' organizing centre induced in the forebrain. Because *Fgf8* itself is expressed in the isthmus, our results identify FGF8 as an important signalling molecule in normal midbrain development.

Fgf8 is expressed in the isthmus of the developing mouse brain^{5–8}. Because FGF8 has inducing activity in another developmental system (the limb⁹), we sought to determine whether FGF8 provides the midbrain-inducing activity of an isthmus graft in the chick. We first confirmed that *Fgf8* is expressed in the chick isthmus (Fig. 1a). Next, we determined the effects of implanting a bead soaked in recombinant FGF8 (FGF8-bead) into the caudal diencephalon (prosomer 2, p2, as defined in ref. 10; Fig. 1b) of chick embryos at stages 9–12 (ref. 11). An early effect of isthmus grafts is induction in the host neuroepithelium of *Engrailed-2* (*En2*) expression¹⁴, an early marker of mes/rhombencephalic development^{12–14}. When an FGF8-bead was implanted, ectopic *En2* RNA was detected caudal to the zona limitans intrathalamica (ZL), a transverse boundary separating dorsal and ventral thalamus anlagen (p2/p3 boundary¹⁰), in all embryos assayed 22–26 h later ($n = 10$; Fig. 1c). Control beads soaked in phosphate-buffered saline (PBS-beads) did not induce *En2* expression ($n = 11$; not shown).

In experimental embryos surviving to E5–E16 (stages 25–42) the diencephalon caudal to the ZL (p1 and p2; ref. 10) was transformed from its normal fate of rostral pretectum and dorsal thalamus to ectopic midbrain ($n = 17/18$; Fig. 2). Control embryos implanted with PBS-beads that survived to E4–E10

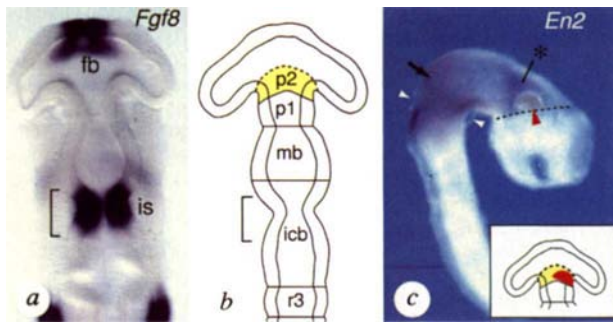


FIG. 1 *Fgf8* is expressed in the isthmus and FGF8 protein induces *En2* gene expression in the caudal diencephalon. **a**, *Fgf8* RNA is detected in a stage 11 chick embryo (dorsal view) in neuroepithelial cells spanning the isthmus (is) (except in the ventral midline; not shown), in cells surrounding the anterior neuropore in the forebrain (fb), in the branchial arch region (bottom of field), and the primitive streak (not shown). **b**, Fate map of the chick brain at stage 11 showing prosomere 2 (p2; highlighted in yellow).

displayed no such abnormalities ($n = 11$; data not shown). Although the ectopic midbrain always formed on the FGF8-bead-implanted side, in some cases p1 on the contralateral side was affected (not shown). Figure 2a, b shows both normal and experimental embryos. The ectopic midbrains appeared either fused with or separate from their normal counterparts (Fig. 2c–f). Sectioning revealed that induced midbrains were always present in a mirror image orientation relative to the normal midbrain (Fig. 2g–i), and contained an isthmus at their rostral end, as shown by the presence of isthmus nuclei. In cases in which the ectopic and normal midbrains were fused, the most rostral midbrain element, the griseum tectalis, was absent from both, and a single, bicaudally symmetric tectal structure was observed (Fig. 2d, g, and data not shown). Respecification of the caudal diencephalon into the midbrain always involved the alar plate, but in two of seven cases examined histologically the basal plate was also clearly transformed, and contained oculomotor, III and IV cranial nerve nuclei, present as mirror image duplications of normal midbrain and isthmus basal plate derivatives (Fig. 2h, i). In all cases the ventral thalamus rostral to the ZL appeared normal (not shown).

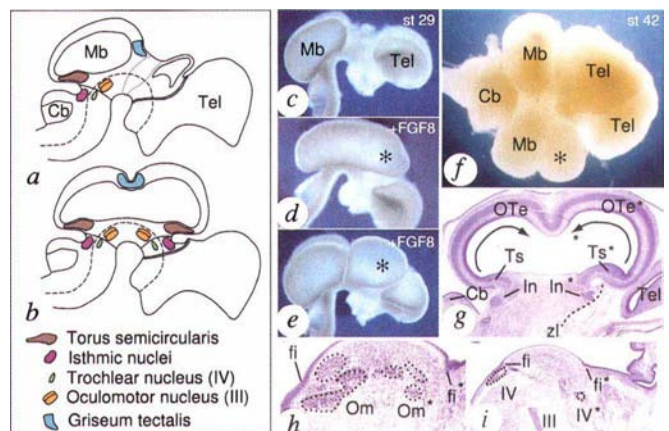
Neither isthmus grafts nor FGF8-beads placed in p2 induce host tissue to form cerebellum (ref. 4, and data not shown). In

contrast, isthmus grafts induce a cerebellar fate in the rhombencephalon¹⁵. However, in preliminary experiments we found that FGF8-beads implanted in the rhombencephalon did not appear to induce ectopic *En2* expression or cerebellar differentiation in embryos assayed at stages 15–19 ($n = 5$) or stages 30–36 ($n = 9$), respectively (data not shown) suggesting that FGF8 alone cannot induce cerebellar tissue in the rhombencephalon.

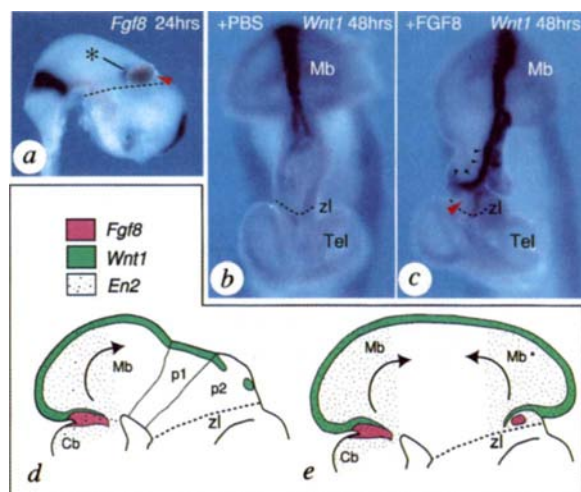
We next determined whether FGF8 induces expression of two genes (in addition to *En2*) normally expressed in the isthmus, *Fgf8* and *Wnt1*. At 22–26 h after bead implantation, *Fgf8* RNA was detected in the vicinity of an FGF8-bead ($n = 10/12$; Fig. 3a), but not a PBS-bead ($n = 6$; data not shown). As observed for *En2*, ectopic *Fgf8* expression was detected only in cells caudal of the ZL (Fig. 3a). *Fgf8* RNA was also detected 48 h after bead implantation ($n = 3/3$; data not shown). These results parallel our recent finding that FGF8 protein induces *Fgf8* gene expression during limb development⁹. From stage 16 onwards, expression of *Wnt1*, whose activity is required for formation of the mid/hindbrain region in the mouse^{16,17}, is normally detected in a transverse ring (excluding the ventral midline), just rostral to the ring of *Fgf8* expression in the isthmus, and also along the dorsal midline (see Fig. 3b, d, and refs 18 and 19). It was difficult to determine whether

FIG. 2 FGF8 induces transformation of the caudal diencephalon to midbrain. **a**, **b**, Schematic representations of brains from normal (**a**); based on Fig. 5 of ref. 2) and experimental (**b**) embryos at stage 36 (E10). The dashed line (longitudinal axis of the brain) separates the alar and basal plates. All structures duplicated in the experimental embryos are illustrated in **b** (individual embryos did not necessarily contain all structures). **c**, Brain from normal embryo at stage 29 (lateral view). **d**, **e**, Brains from experimental embryos at stage 29 showing regions transformed from diencephalic to mesencephalic fate (*). In **d**, the ectopic midbrain is fused with its normal counterpart; in **e** the ectopic midbrain is complete and separate from the normal midbrain. **f**, Brain from experimental embryo at stage 42 (dorsal view) showing ectopic midbrain (*). Note that unoperated side (top) appears unaffected. **g**, Brain of experimental embryo at stage 35 (sagittal section) showing transformation of caudal diencephalon to isthmus and midbrain structures (*). Arrows indicate polarity, from caudal to rostral, of normal and ectopic midbrains. Note the reversed polarity of the ectopic midbrain structures and isthmus nuclei juxtaposed to the zona limitans. **h**, **i**, More medial sections of brain shown in **g**, illustrating mirror image duplication of motor nuclei and nerves in basal plate. Cb, Cerebellum; fi, fovea isthmi; In, isthmus nuclei; Mb, midbrain; Om, oculomotor nuclei; OTe, optic tectum; Tel, telencephalon; Ts, torus semicircularis; zli, zona limitans; III, third cranial nerve; IV, fourth cranial nerve.

METHODS. Heads from experimental embryos were fixed in 4% paraformaldehyde, washed in PBS, 0.1% Tween 20, and dissected free of



surrounding tissue before photography. Tissue for histological analysis was fixed overnight in 75% ethanol, 25% glacial acetic acid, embedded in wax, sectioned at 10 μ m, and stained with cresyl violet.



Wnt1 expression was altered in embryos assayed 24 h after bead implantation, because at that stage the rostral end of the normal *Wnt1* expression domain still lies close to the site of bead insertion. However, 48 h after bead placement, the *Wnt1* domain clearly extended rostral to its normal anterior limit and also laterally around the FGF8-bead ($n = 5/5$; Fig. 3c). No ectopic *Wnt1* expression was detected when PBS-beads were implanted ($n = 9$; Fig. 3b). Thus, during the 48 h after FGF8-bead placement, three genes normally expressed in the isthmus, *En2*, *Fgf8*, and *Wnt1*, are induced in the diencephalon in the vicinity of the bead (summarized in Fig. 3e).

These data demonstrate that FGF8 protein is sufficient to induce a complete midbrain in competent neuroepithelium and suggest that FGF8 induces the ectopic structures by establishing an isthmus-like organizing centre in the caudal diencephalon. Similar results were obtained with beads soaked in FGF4 protein (not shown), presumably reflecting the ability of FGF4 to activate the same signal transduction pathway as does FGF8. Signals from the induced organizing centre in p2 appear to have a long-range effect, because they transform both p2 and p1 to a midbrain fate. The reversed polarity of the ectopic midbrain indicates that tissue closest to the organizing centre is transformed to a caudal fate, whereas tissue further away is respecified to a more rostral fate. In some cases the most rostral structure, the griseum tectalis, is absent from both the normal and ectopic midbrains. An analogous effect is seen in limb buds with polarizing regions on opposite sides, which sometimes develop a mirror-image digit pattern (4-3-3-4) that lacks the most anterior digit (digit 2)²⁰. These effects are presumably due to overlap of patterning activities emanating from

FIG. 3 FGF8-beads induce ectopic *Fgf8* and *Wnt1* expression in the caudal diencephalon. **a**, Expression of *Fgf8* in experimental embryo at stage 16, 24 h after bead implantation (lateral view), showing domain of ectopic *Fgf8* expression (*) caudal to the zona limitans (dashed line) in vicinity of FGF8-bead (red arrowhead). **b**, **c**, *Wnt1* gene expression in control (**b**) and experimental (**c**) embryos at stage 19, 48 h after bead implantation (dorsal view) showing ectopic expression of *Wnt1* (small arrowheads), which is localized in the caudal diencephalon both rostral and lateral to the normal *Wnt1* expression domain. Red arrowhead points to FGF8-bead, and dashed line marks zona limitans. **d**, **e**, Schematic diagrams illustrating domains of gene expression at stage 19 in the midbrain and caudal diencephalon of normal (**d**) and experimental (**e**) embryos. *Fgf8* and *Wnt1* expression patterns are based on observations made 48 h after bead implantation, whereas *En2* expression pattern is inferred from observations made after 24 h. Arrows indicate caudal to rostral polarity of midbrains. Abbreviations as in Figs 1 and 2.

METHODS. FGF8- or PBS-beads were inserted into p2 at stage 9–12, and embryos were fixed after 24 or 48 h incubation. *Fgf8* expression was detected as described in legend to Fig. 1. *Wnt1* expression was detected by whole-mount RNA *in situ* hybridization using a chick *Wnt1* antisense probe¹⁹ kindly provided by M. Wassef (Paris).

opposite ends of the responding tissue. FGF-beads appear to affect only tissue caudal to the ZL, perhaps reflecting a lack of competence of the neuroepithelium rostral to the ZL to respond to the inductive effects of FGF; alternatively, the ZL might act as a barrier that blocks the rostral transmission of these effects, consistent with the idea that preexisting boundaries are important in restricting the spread of morphogenetic signals during the ongoing patterning of the neural tube¹⁵.

Midbrain development requires *Engrailed* (*En*) activity^{21,22}. There is evidence that expression of *En1* and *En2* is induced in the neuroectoderm by a signal from the underlying mesoderm²³ and then maintained by a *Wnt1*-dependent signal acting within the plane of the neuroepithelium²⁴. The data reported here demonstrating that FGF8 is sufficient to induce *En* expression (at least *En2*), together with the observation that *Fgf8* is expressed in mesoderm (prospective cardiogenic cells) underlying the neural plate at a stage before *En* induction^{7,8}, raises the possibility that FGF8 produced in this mesoderm functions normally to induce the expression in the isthmus region of *En* and perhaps other genes. A second possible role for FGF8 in normal midbrain development is an involvement in the *Wnt1*-dependent maintenance of *En* gene expression, as suggested by observations that the *Fgf8* expression domain is immediately caudal to the *Wnt1* expression domain in the isthmus region⁸, and products of these two genes cooperate to accelerate mammary tumour formation in mice²⁵. Interactions between WNT1 and FGF8 may be responsible for maintaining not only *En* expression, but perhaps for other aspects of the organizing activity of the isthmus region during normal development. □

Received 4 December 1995; accepted 4 January 1996.

- Bally-Cuif, L., Alvarado-Mallart, R. M., Darnell, D. K. & Wassef, M. *Development* **115**, 999–1009 (1992).
- Marin, F. & Puellas, L. *Dev. Biol.* **163**, 19–37 (1994).
- Joyner, A. L. *Trends Genet.* **12**, 15–20 (1996).
- Martinez, S., Wassef, M. & Alvarado-Mallart, R. M. *Neuron* **6**, 971–981 (1991).
- Heikinheimo, M., Lawshé, A., Shackleford, G. M., Wilson, D. B. & MacArthur, C. A. *Mech. Dev.* **48**, 129–138 (1994).
- Ohuchi, H. et al. *Biochem. biophys. Res. Commun.* **204**, 882–888 (1994).
- Mahmood, R. et al. *Curr. Biol.* **5**, 797–806 (1995).
- Crossley, P. H. & Martin, G. R. *Development* **121**, 439–451 (1995).
- Crossley, P. H., Minowada, G., MacArthur, C. A. & Martin, G. R. *Cell* **84**, 127–136 (1996).
- Rubenstein, J. et al. *Science* **266**, 578–580 (1995).
- Hamburger, V. & Hamilton, H. J. *Morphol.* **88**, 49–92 (1951).
- Davis, C. A. & Joyner, A. L. *Genes Dev.* **2**, 1736–1744 (1988).
- Davis, C. A., Noble-Topham, S., Rossant, J. & Joyner, A. L. *Genes Dev.* **2**, 361–371 (1988).
- Gardner, C. A. & Barald, K. F. *Dev. Dynamics* **193**, 370–388 (1992).
- Martinez, S., Marin, F., Nieto, M. A. & Puellas, L. *Mech. Dev.* **51**, 289–303 (1995).
- McMahon, A. P. & Bradley, A. *Cell* **62**, 1073–1085 (1990).

- Thomas, K. R. & Capocchi, M. R. *Nature* **346**, 847–850 (1990).
- Hollyday, M., McMahon, J. & McMahon, A. *Mech. Dev.* **52**, 9–25 (1995).
- Bally-Cuif, L. & Wassef, M. *Development* **120**, 3379–3394 (1994).
- Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199–202 (1975).
- Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A. B. & Joyner, A. L. *Science* **269**, 679–682 (1995).
- Wurst, W., Auerbach, A. & Joyner, A. *Development* **120**, 2065–2075 (1994).
- Ang, S.-L. & Rossant, J. *Development* **118**, 139–149 (1993).
- McMahon, A., Joyner, A., Bradley, A. & McMahon, J. *Cell* **69**, 581–595 (1992).
- MacArthur, C. A., Shankar, D. B. & Shackleford, G. M. *J. Virol.* **69**, 2501–2507 (1995).
- Marin, F. & Puellas, L. *Eur. J. Neurosci.* **7**, 1714–1738 (1995).
- Wilkinson, D. G. *In situ Hybridisation: A Practical Approach* (IRL at Oxford University Press, Oxford, 1992).
- Bally-Cuif, L. & Wassef, M. *Curr. Opin. Genet. Dev.* **5**, 450–458 (1995).

ACKNOWLEDGEMENTS. We thank C. MacArthur for providing the FGF8 protein used in this study, and to our laboratory colleagues for critical reading of the manuscript. This work was supported by the NIH/LB Program of Excellence in Molecular Biology (G.R.M.), and grants from the NIH and Human Frontiers of Science Program (S.M.). P.H.C. was supported by a postdoctoral fellowship from the Human Frontiers of Science Program.

Dissociation in prefrontal cortex of affective and attentional shifts

R. Dias, T. W. Robbins & A. C. Roberts*†

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, UK

The prefrontal cortex is implicated in such human characteristics as volition, planning, abstract reasoning and affect¹⁻⁶. Frontal-lobe damage can cause disinhibition such that the behaviour of a subject is guided by previously acquired responses that are inappropriate to the current situation⁷⁻⁹. Here we demonstrate that disinhibition, or a loss of inhibitory control, can be selective for particular cognitive functions and that different regions of the prefrontal cortex provide inhibitory control in different aspects of cognitive processing. Thus, whereas damage to the lateral prefrontal cortex (Brodmann's area 9) in monkeys causes a loss of inhibitory control in attentional selection, damage to the orbitofrontal cortex in monkeys causes a loss of inhibitory control in 'affective' processing, thereby impairing the ability to alter behaviour in response to fluctuations in the emotional significance of stimuli. These findings not only support the view that the prefrontal cortex has multiple functions, but also provide evidence for the distribution of different cognitive functions within specific regions of prefrontal cortex.

The apparent heterogeneity of functional deficits, and the variability of lesions resulting from frontal-lobe damage in man has made it difficult to assign specific functions to particular prefrontal regions^{2,6}. But it is possible that this heterogeneity reflects impairments in a common set of cognitive functions operating across different sensory modalities (for example, visual versus spatial). There is evidence from studies in non-human primates for a global role of the prefrontal cortex in a process that holds representations of stimulus information 'on-line': an important component of 'working memory'³, with independent analysis of visual and spatial information in adjacent regions¹⁰. Our study provides evidence for another general process, response inhibition, that operates within different regions of the prefrontal cortex to affect different forms of cognitive processing, even within the same modality, thus promoting cognitive flexibility. We devised a procedure for examining dissociable forms of inhibitory control of behaviour that operate at two different levels of response selection, namely changes in assignment of affective association to specific visual stimuli, and shifts in selective attention from one perceptual dimension of a complex visual stimulus (such as its form) to another¹¹.

Nine marmoset monkeys were trained to make visual discriminations between two compound stimuli, each consisting of a black line superimposed on a blue polygon (Fig. 1a). Selection of the correct stimulus led to a reward, and monkeys learned to identify the stimulus by trial and error. For some monkeys, the correct response depended on the shape of the black line and for other monkeys on the shape of the blue polygon. Monkeys were trained to maintain an attentional set (the tendency to respond to a particular perceptual dimension on the basis of previous experience) towards one dimension (either the polygons or the lines) over a series of discriminations with novel stimuli. The shape of the other dimension was uncorrelated with the reward. The monkeys then received infusions of the excitotoxin, quinolinic acid, into either the orbital (ORB group $n = 3$) or lateral (LAT group, $n = 3$) regions of prefrontal cortex, or

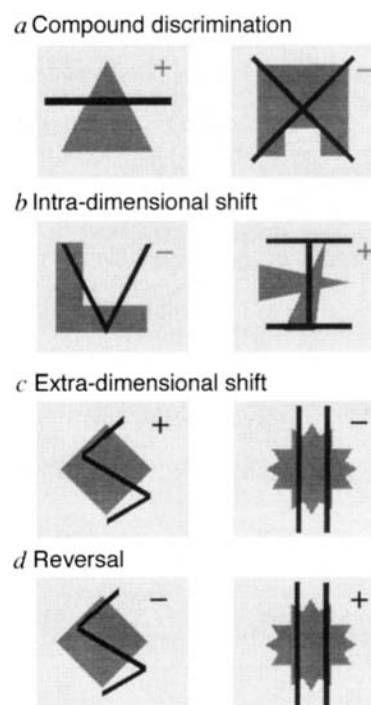


FIG. 1 *a-d*, Examples of the compound stimuli used in various stages of the attentional set-shifting procedure. They are identical to those used on a computerized touch-screen version of the test except that in the latter, the lines are white, rather than black. In these examples the dimension of 'blue polygon' is relevant in all discriminations except those requiring an extra-dimensional shift (*c*) and reversal (*d*). In any one trial of a compound discrimination a polygon exemplar may be paired with one or other line exemplar. Correct and incorrect choices are indicated by plus (+) and minus (-), respectively. Grey typeface specifies that 'polygon' is the relevant dimension and black typeface specifies that 'line' is the relevant dimension.

METHODS. The monkeys sat in a Perspex transport cage and on removal of one side they were able to look through a window with Perspex bars. On the other side of the window were two wooden test boxes that were open on only one side, positioned on the far right and far left of a shelf. Attached to the front of each test box was a piece of clear Perspex behind which the stimulus cards were placed. Two screens, one clear and one opaque, were positioned between the window and test boxes while each trial was prepared. Removal of the opaque screen indicated the beginning of each trial. Tapping the screen in front of one of the stimulus cards resulted in its removal, enabling the monkey to turn the test box around and retrieve the reward, a small piece of marshmallow, from within. A tap in front of the other stimulus card resulted in the opaque screen being placed in front of the test boxes. For a full description of the behavioural sequence, see ref. 16.

sham surgery (control group, $n = 3$). Subsequent histological analysis showed that the injections of excitotoxin destroyed selectively the vast majority of neurons in the infused regions (Figs 2, 3).

After surgery, the ability of monkeys to remember the discrimination they had learnt immediately before surgery was assessed (retention test). There were no differences between the three groups. Their ability to learn a series of three compound discriminations was then assessed using novel stimuli in which the previously relevant dimension remained relevant (intra-dimensional shifts (Fig. 1*b*)). Again, there were no significant differences between the groups, demonstrating that all monkeys were able to maintain an attentional set and transfer behavioural control from one pair of exemplars to another within the same perceptual dimension. Significant differences between the groups emerged

* To whom correspondence should be addressed at the Department of Anatomy.

† Present address: Dept of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK.

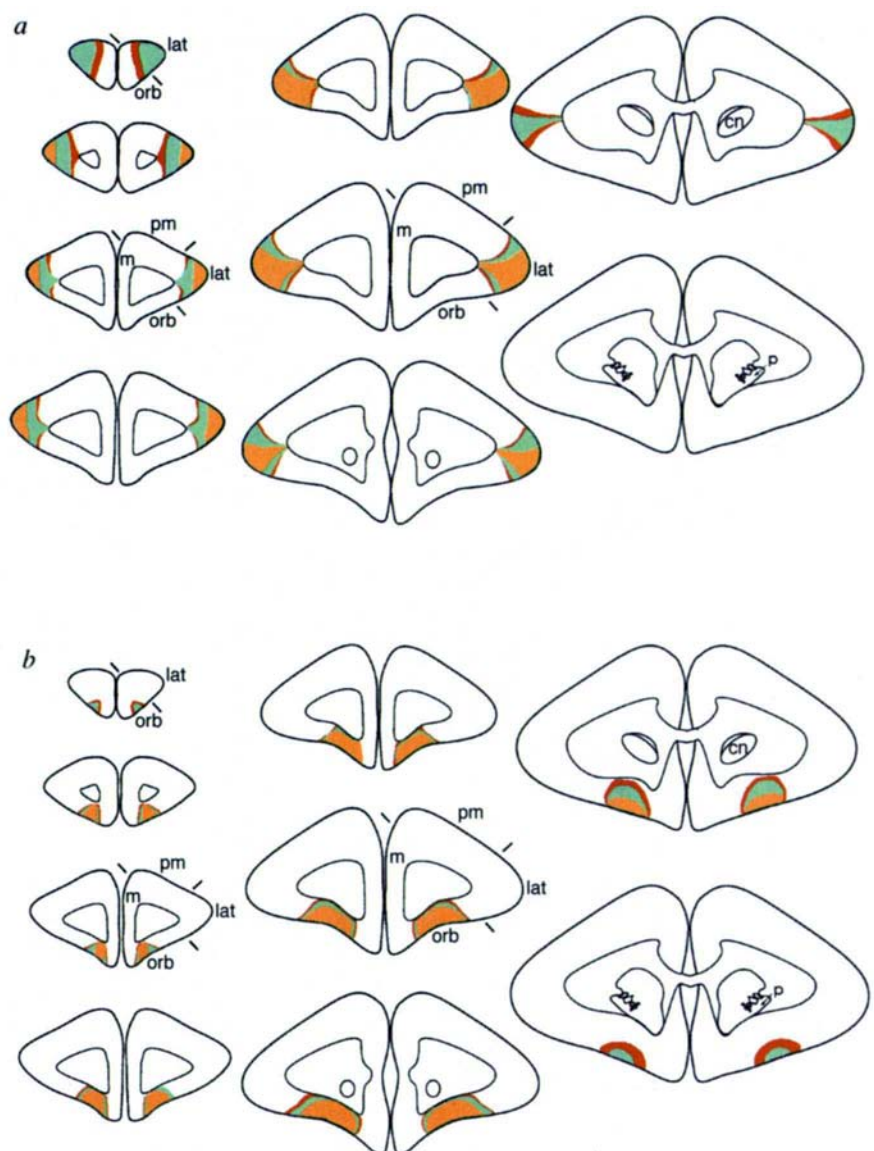
when they had to learn a novel compound discrimination that required a shift of attentional set from one dimension to another (extra-dimensional shift) (Fig. 1c). Lesions of lateral prefrontal cortex, but not orbitofrontal cortex, impaired attentional set-shifting ability, with monkeys with lesions of the lateral prefrontal cortex taking twice as many trials to reach a particular level of accuracy of performance as either monkeys with lesions of orbitofrontal cortex or monkeys that received sham surgery (Fig. 4). In contrast, reversing a stimulus–reward association between the same pair of exemplars within a dimension (Fig. 1d), as distinct from shifting an attentional set between dimensions, was impaired by a lesion of the orbitofrontal cortex and not the lateral prefrontal cortex (Fig. 4). This deficit involves a failure to suppress the influence of previously acquired stimulus–reward associations, rather than an impairment in learning new stimulus–reward associations (as is required for an intra-dimensional shift).

This double dissociation between the behavioural effects of lateral and orbitofrontal cortex lesions in the marmoset provides

new insight into the functional organization of the prefrontal cortex in primates. These results cannot be incorporated easily into an account of prefrontal function based purely on holding information ‘on-line’ in working memory, as whereas each of the discriminations requires the short term storage of information, only those involving inhibition of a previously acquired response were impaired. Moreover, using a modified version of the same discrimination paradigm, it has been shown that frontal lobe damage in humans impairs set-shifting ability only when the previously relevant dimension is still present and not when it is replaced by a new dimension¹². This provides further evidence against an account of prefrontal functioning that emphasises only the retention of information over short intervals, as the working memory load was equated across these two types of attentional shift.

Our findings suggest that distinct regions of the prefrontal cortex carry out independent, but complementary, forms of cognitive processing of complex visual stimuli in changing environmental circumstances. These processes may be organized in an

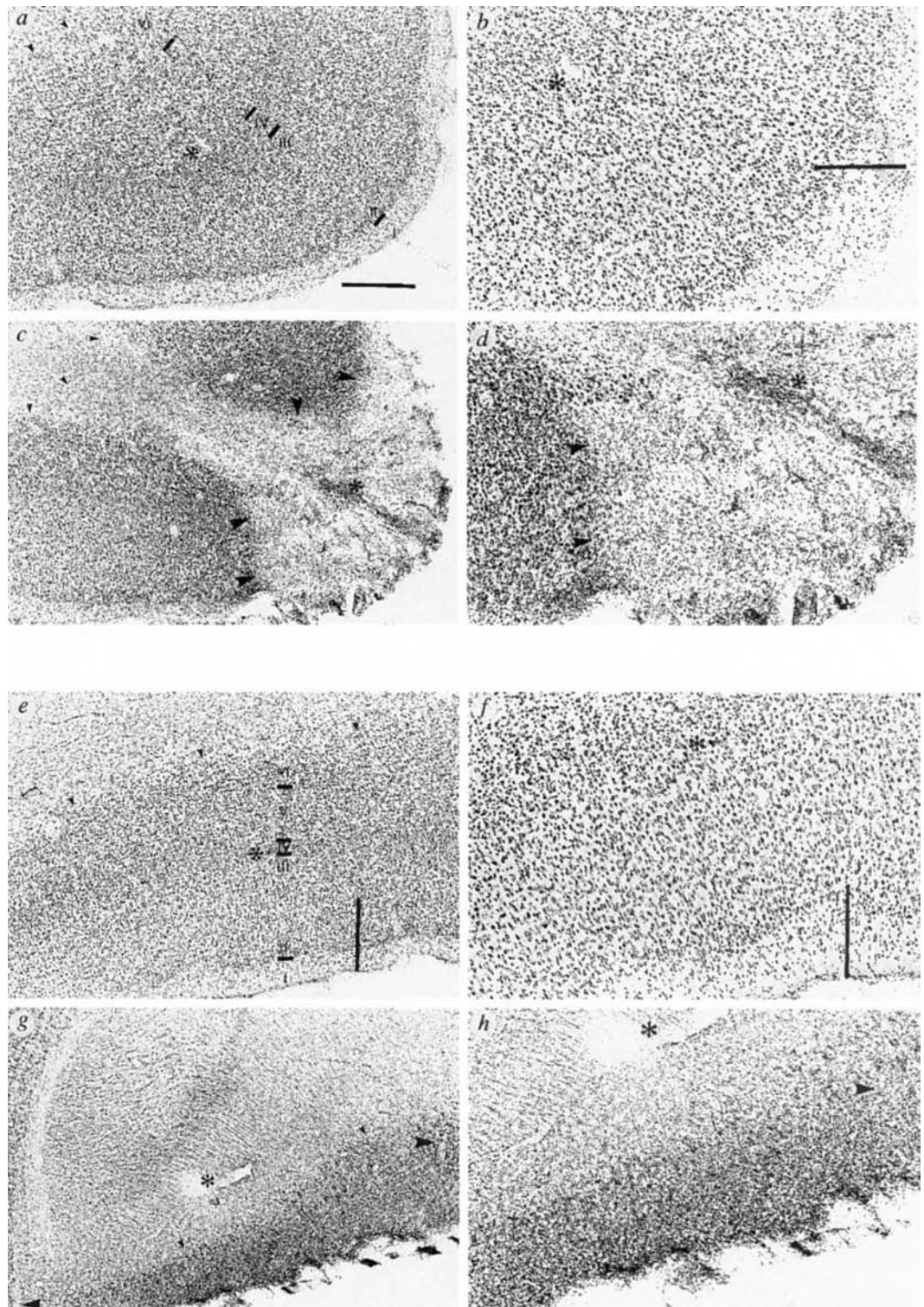
FIG. 2 A series of coronal sections through the prefrontal cortex illustrating the lateral (a) and orbital (b) prefrontal lesions for each of the three marmosets (pale, medium and dark stippling). Lateral (lat) prefrontal cortex corresponds to Brodmann's area 9 (ref. 17) and comprises a thick, well-developed layer IV with small pyramidal neurons within its deeper layers. Layer IV in adjacent cortex is less well-defined, making the boundaries to this region easily observable. On the basis of criteria used to compare the cytoarchitectonics of *Macaca* and *Galago*¹⁸, this lateral region of cortex in the marmoset is probably homologous to Walker's areas 9, 10 or 12 of *Macaca*. For a definition of orbital prefrontal cortex, see ref. 18. There is no overlap between the two lesioned areas. In the LAT group (a) cell loss extends from the frontal pole, anteriorly, to the level of the rostral limit of the caudate nucleus (cn), posteriorly, being most variable within the frontal pole (first 2–3 sections). Cortex in the adjacent premotor areas (pm) and orbital areas (orb) is spared in the LAT group. In the ORB group (b) the most anterior limit of cell loss is at the level of the appearance of the white matter whereas the most posterior limit extends just rostral to the putamen (p). Cortex is spared in the adjacent lateral (lat) and medial (m) prefrontal regions. METHODS. Marmosets were anaesthetized with pentobarbitone (0.15 ml of a 60 mg solution, intra peritoneally) and placed in a stereotaxic frame. The lateral and orbitofrontal cortex was destroyed by injecting 0.6–1.5 μ l per site of a 0.09 M solution of quinolinic acid (in 0.01 M phosphate buffer, pH 7.0) bilaterally into several sites within the respective regions of cortex via a stainless steel cannula (30 gauge).



hierarchical manner¹³ or in parallel, but each depends on a process of response inhibition. Thus, the capacity to reverse affective associations for specific visual stimuli is mediated by regions within the orbitofrontal cortex, consistent with its close connections with the limbic system^{14,15}. In contrast, regions within the

lateral prefrontal cortex may be implicated in the higher-order shifting of attention between supra-ordinate features of visual stimuli, such as their perceptual dimensions. These findings in a New World monkey pose questions about how similar processes may be organized anatomically within the prefrontal cortex of

FIG. 3 Photomicrographs of representative cresyl-violet-stained coronal sections through an intermediate level of the frontal pole depicting lateral prefrontal cortex in an intact (*a, b*) and lateral cortex-lesioned brain (*c, d*) at low (*a, c*) and high power (*b, d*) and orbitofrontal cortex in an intact (*e, f*) and orbitofrontal cortex-lesioned brain (*g, h*) at low (*e, g*) and high power (*f, h*). Small arrows mark the boundary between the cortex and the white matter. The dorsal and medial extent of cell loss after a lesion of the lateral prefrontal cortex can be readily seen in *c* (large arrows mark the boundaries). Enlargement of the area in (*d*) shows the almost total loss of all 6 layers of neurons in the centre of the lesioned area, in contrast to the dense layering of neurons seen in the intact brain (*b*). The medial and lateral boundaries of cell loss after a lesion of the orbitofrontal cortex are depicted by the large arrows in *g*. There is far greater gliosis within this region of the frontal pole than is seen in the lateral region. Again, the almost total loss of neurons within the lesioned area (*h*) is in stark contrast to the dense layering of neurons present in the intact brain (*f*). By comparing the width of the tissue on the orbital surface of the lesioned brain in *g* and the intact brain in *e*, the cortex in the lesioned brain can be seen to be about a quarter of the thickness of the intact brain. Asterisks mark the same locations in each pair of low and high power photomicrographs. Scale bars in *a* and *e* represent 500 μm ; in *b* and *f*, 300 μm .



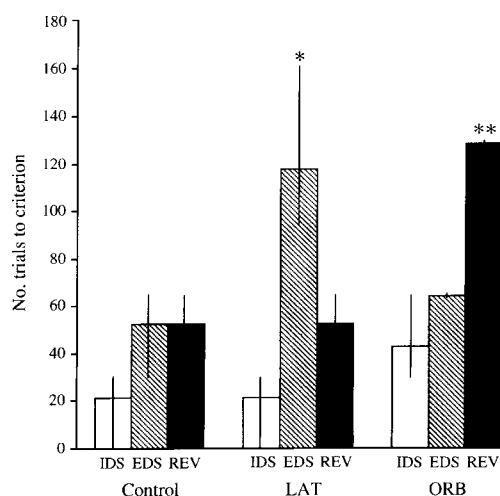


FIG. 4 Number of trials taken to reach a criterion level of performance in a visual discrimination task requiring an intra-dimensional shift, IDS (open bars) (final discrimination of the series of three presented post-operatively), an extra-dimensional shift, EDS (hatched bars) and a reversal, REV (black bars) in monkeys that received an excitotoxic lesion of either the lateral (LAT) or orbitofrontal cortex (ORB) or a sham operation (Control). Bars on the histogram represent the range of scores of the monkeys within each group. Whereas lesions of lateral prefrontal cortex increased selectively the number of trials made before attaining criterion on the EDS but not the IDS or reversal, lesions of the orbitofrontal cortex increased selectively the number of trials made before attaining criterion on the reversal but not the IDS or the EDS. An analysis of trials to criterion using the factors 'lesion' (lateral, orbital and control) and 'test' (IDS, EDS and reversal) showed a significant interaction of lesion $\times$ test ($F_{4,12} = 10.353$, $P < 0.001$). Post-hoc analysis of the simple main effects showed that there was a profound effect of lesioning on the EDS ($F_{2,17} = 9.3$, $P < 0.01$) and on the reversal ($F_{2,17} = 14.7$, $P < 0.001$) but not on the IDS ($F > 1$). Monkeys with lesions of the orbitofrontal cortex made significantly more perseverative responses on the reversal, that is responses to the exemplar that had been rewarded previously, than either controls or monkeys with lesions of the lateral prefrontal cortex ($F_{2,6} = 7.372$, $P < 0.05$). (All errors within each half session (16 trials) were defined as perseverative if the animal's performance across the 16 trials was significantly below chance). * The LAT group was significantly different to the ORB and control groups at the 1% level. ** The ORB group was significantly different to the LAT and control groups at the 1% level. (The same pattern of effects was seen with an ANOVA test on the error score.)

higher primates, including man, and how these functions might be integrated with 'working-memory' processes to provide executive control of behaviour. □

Received 8 June; accepted 21 December 1995.

1. Fuster, J. M. *The Prefrontal Cortex* 2nd edn (Raven, New York, 1989).
2. Damasio, A. R. *Descartes' Error. Emotion, Reason and the Human Brain* (Putnam, New York, 1994).
3. Goldman-Rakic, P. S. in *Handbook of Physiology: the Nervous System*, Vol. 5 (ed. Plum, F. & Mountcastle V.) 373–417 (American Physiological Society, Bethesda, 1987).
4. Passingham, R. *The Frontal Lobes and Voluntary Action* (Oxford Univ. Press, Oxford, 1993).
5. Petrides, M. *Proc. R. Soc. Lond. B* **246**, 293–298 (1991).
6. Shallice, T. in *From Neuropsychology to Mental Structure* 328–350 (Cambridge Univ. Press, 1988).
7. Milner, B. *Archs Neurol.* **9**, 100–110 (1963).
8. Mishkin, M. in *The Frontal Granular Cortex and Behaviour* (ed. Warren, J. M. & Akert, K.) 219–294 (McGraw-Hill, New York, 1964).
9. Iversen, S. D. & Mishkin, M. *Expl Brain Res.* **11**, 376–386 (1970).
10. Wilson, F. A. W., O'Scalaidhe, S. P. & Goldman-Rakic, P. S. *Science* **260**, 1955–1957 (1993).
11. Roberts, A. C., Robbins, T. W. & Everitt, B. J. *Q. J. exp. Psychol.* **40**, 321–341 (1988).
12. Owen, A. M. et al. *Brain* **116**, 1159–1175 (1993).
13. Sutherland, N. S. & Mackintosh, N. J. *Mechanisms of Animal Discrimination Learning* (Academic, New York, 1971).
14. Nauta, W. J. H. *J. Psychiatr. Res.* **8**, 167–187 (1971).
15. Amaral, D. G., Price, J. L., Pitkanen, A. & Carmichael, S. T. in *The Amygdala* (ed. Aggleton, J. P.) 1–66 (Wiley-Liss, New York, 1993).
16. Roberts, A. C. et al. *J. Neuroscience* **14**, 2531–2544 (1994).
17. Brodmann, K. *Vergleichende Lokalisationslehre der Großhirnrinde* (Barth, Leipzig 1909).
18. Preuss, T. M. & Goldman-Rakic, P. S. *J. comp. Neurol.* **310**, 429–474 (1991).

ACKNOWLEDGEMENTS. This work was supported by a University Research Fellowship grant from the Royal Society (A.C.R.) and by a Programme grant from the Wellcome Trust. R.D. received an MRC studentship. We thank R. M. Ridley for supplying the marmosets, J. Bashford and I. Bolton for photographic assistance, and B. J. Everitt for critical reading of the manuscript and preparation of the photomicrographs.

Enhanced dihydropyridine receptor channel activity in the presence of ryanodine receptor

Junichi Nakai*†‡, Robert T. Dirksen†, Hanh T. Nguyen§, Isaac N. Pessah||, Kurt G. Beam† & Paul D. Allen§¶#

* Department of Medical Chemistry, Kyoto University Faculty of Medicine, Kyoto 606-01, Japan

† Department of Anatomy and Neurobiology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado 80523, USA

§ Department of Cardiology, Laboratory of Molecular Cardiology, Children's Hospital, Boston, Massachusetts 02115, USA

|| Department of Molecular Biosciences, University of California-Davis, School of Veterinary Medicine, Davis, California 95616, USA

¶ Department of Anesthesiology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA

EXCITATION-CONTRACTION coupling in skeletal muscle involves a voltage sensor in the plasma membrane which, in response to depolarization, causes an intracellular calcium-release channel to open. The skeletal isoform of the ryanodine receptor (RyR-1) functions as the Ca^{2+} -release channel^{1–3} and the dihydropyridine receptor (DHPR) functions as the voltage sensor and also as an L-type Ca^{2+} channel^{4,5}. Here we examine the possibility that there is a retrograde signal from RyR-1 to the DHPR, using myotubes from mice homozygous for a disrupted RyR-1 gene (dyspedic mice)³. As expected, we find that there is no excitation-contraction coupling in dyspedic myotubes, but we also find that they have a roughly 30-fold reduction in L-type Ca^{2+} -current density. Injection of dyspedic myotubes with RyR-1 complementary DNA restores excitation-contraction coupling and causes the density of L-type Ca^{2+} current to rise towards normal. Despite the differences in Ca^{2+} -current magnitude, measurements of charge movement indicate that the density of DHPRs is similar in dyspedic and RyR-1-expressing myotubes. Our results support the possibility of a retrograde signal by which RyR-1 enhances the function of DHPRs as Ca^{2+} channels.

Intracellular Ca^{2+} responses produced by application of caffeine and ryanodine to normal myotubes, dyspedic myotubes and dyspedic myotubes injected with RyR-1 cDNA were measured (Fig. 1). In a normal myotube (Fig. 1a), application of 10 mM caffeine caused intracellular Ca^{2+} release, and the subsequent addition of 0.1 mM ryanodine caused a gradual elevation in Ca^{2+} accompanied by slow Ca^{2+} oscillations. A second application of caffeine after removal of ryanodine caused a smaller transient release than the initial application, possibly because micromolar concentrations of ryanodine cause opening of the ryanodine receptor channel (tending to deplete Ca^{2+} stores) followed by blocking of the channel⁶. Dyspedic myotubes often behaved qualitatively like normal myotubes, but higher caffeine concentrations were required to elicit transient release of Ca^{2+} (Fig. 1b) and the responses were substantially smaller. Attenuated caffeine responses have previously been reported for dyspedic muscle^{3,7} and have been attributed to the presence of the brain isoform

‡ Present Address: Department of Anatomy and Neurobiology, Colorado State University, Fort Collins, Colorado 80523, USA.

To whom correspondence should be addressed.

TABLE 1 Charge movement and Ca^{2+} conductance in normal, dysgenic, dyspedic and RyR-1-expressing dyspedic myotubes

Myotubes	Q_{max} (nC/ μF)	$\bar{V}_0$ (mV)	k_0 (mV)	G_{max} (nS/nF)	$\bar{V}_G$ (mV)	k_G (mV)	$G_{\text{max}}/Q'_{\text{max}}$ (nS/pC)
Normal	5.2 ± 2.6 (23)	-1.2 ± 9.2 (23)	13.2 ± 2.1 (23)	333.4 ± 68.5 (14)	11.5 ± 10.1 (14)	4.9 ± 0.8 (14)	111.1
Dysgenic	2.2 ± 0.4 (9)	-14.4 ± 6.4 (9)	11.0 ± 4.8 (9)	—	—	—	—
Dyspedic	4.0 ± 1.4 (10)	-6.4 ± 9.2 (10)	12.6 ± 1.5 (10)	22.4 ± 28.1 (28)	14.6 ± 12.6 (16)	7.2 ± 1.6 (16)	12.4
RyR-1-expressing	4.0 ± 1.5 (12)	-5.2 ± 8.1 (12)	12.4 ± 1.8 (12)	145.3 ± 46.9 (12)	18.4 ± 9.4 (12)	5.5 ± 1.0 (12)	80.7

Data are given as mean $\pm$ s.d., with the number of measurements given in parentheses. Q_{ON} was measured as previously described and fitted according to $Q_{\text{ON}} = Q_{\text{max}} / \{1 + \exp[-(V - \bar{V}_0)/k_0]\}$ (ref. 11). Values of maximal conductance, G_{max} , were obtained by fitting the peak currents according to $I = G_{\text{max}}(V - V_{\text{rev}}) / \{1 + \exp[-(V - \bar{V}_G)/k_G]\}$ (ref. 11). In the calculation of average G_{max} , a value of 0 nS/nF was used for dyspedic myotubes exhibiting ≤ 0.2 pA/pF ($n = 12$). Brackets indicate data sets compared statistically using a two-sample *t*-test; an asterisk indicates a statistically significant difference ($P < 0.01$), whereas no asterisk indicates $P > 0.05$. There were significant ($P < 0.01$) differences in $\bar{V}_0$ for dysgenics compared to normals (compare with ref. 11) and in k_0 for dyspedic myotubes compared with both normal and RyR-1-expressing dyspedic myotubes. None of the other differences were significant ($P > 0.01$) for the parameters of voltage dependence ($\bar{V}_0$, k_0 , $\bar{V}_G$, k_G). Q'_{max} is the difference between Q_{max} and $Q_{\text{dysgenic-max}}$ (2.2 nC/ μF). The estimated series resistance error was <10 mV for all data.

(RyR-3) of the ryanodine receptor^{7,8} in muscle. The observation that ryanodine occasionally caused a slow rise in Ca^{2+} concentration in dyspedic myotubes (Fig. 1b) supports the possibility of involvement of some form of ryanodine receptor in the attenuated caffeine responses. Dyspedic myotubes expressing RyR-1 cDNA were identified by spontaneous or electrically evoked contractions, which were never observed in non-injected dyspedic myotubes. RyR-1-expressing myotubes were sensitive to caffeine (Fig. 1c), often at even lower concentrations than normal myotubes (Fig. 1, legend). This increased sensitivity (compare with ref. 9) may result from an abnormally high local density of ryanodine receptors near sites of cDNA injection.

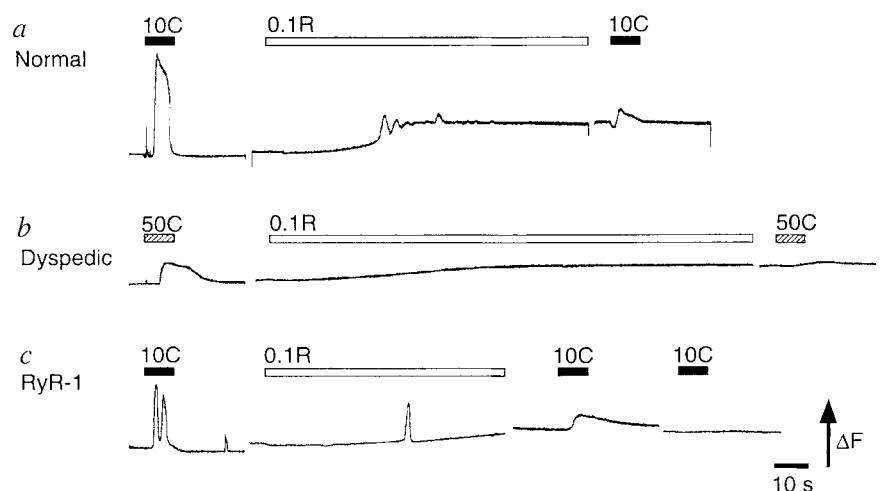
Ca^{2+} transients and whole-cell membrane currents were recorded from normal, dyspedic and RyR-1 expressing myotubes (Fig. 2). Depolarization of normal myotubes elicited slow L-type Ca^{2+} currents and intracellular Ca^{2+} transients (Fig. 2a) similar to

those previously described¹⁰. Moreover, the Ca^{2+} transient was affected little by blocking of the ionic current, consistent with skeletal-type excitation-contraction coupling^{5,10}. By contrast, depolarization elicited little L-type current and failed to evoke a Ca^{2+} transient in dyspedic myotubes (Fig. 2b). Dyspedic myotubes expressing RyR-1 resembled normal myotubes, displaying an appreciable L-type Ca^{2+} current and a depolarization-induced Ca^{2+} transient, which persisted after blockage of the ionic current (Fig. 2c).

A more detailed comparison of L-type Ca^{2+} currents recorded from normal, dyspedic and RyR-1-expressing dyspedic myotubes is presented in Fig. 3. Slowly activating L-type currents were present in each cell type (Fig. 3a–c) with similar voltage dependence (Fig. 3d), but the density of peak current was much lower in the non-injected dyspedic myotubes (0.6 ± 0.7 pA/pF, $n = 28$) than in normal myotubes (16.8 ± 4.3 pA/pF, $n = 14$). Some

FIG. 1 Intracellular Ca^{2+} transients in response to caffeine (C) and ryanodine (R) in normal, dyspedic and RyR-1-expressing dyspedic myotubes. Traces are interrupted between drug applications, and indicated concentrations are in mM. a, Normal myotubes produced Ca^{2+} transients in response to 10 mM caffeine ($n = 11/11$) but failed to respond to 1 mM caffeine ($n = 0/10$). b, Dyspedic myotubes never responded to 1 mM caffeine ($n = 0/9$), but sometimes responded to higher concentrations (10 mM, $n = 13/26$; 50 mM, $n = 5/11$). c, RyR-1-expressing dyspedic myotubes responded to low concentrations of caffeine (1 mM, $n = 11/12$; 10 mM, $n = 4/4$). Whenever there was a strong response to the first caffeine exposure, the application of ryanodine caused a large reduction in the second caffeine response, even if the ryanodine did not cause a gradual increase in Ca^{2+} as shown here. The decrement in caffeine responses was usually less than 20% when there was no intervening ryanodine exposure.

METHODS. The gene-targeting of J1 embryonic-stem cells and creation of germ line $+/\text{ryr1}$ chimaeric mice with one intact and one interrupted RyR-1 gene (H.T.N. *et al.*, manuscript in preparation) was done using standard techniques^{24,25}. Interruption of the RyR-1 gene was confirmed by Southern-blot analysis. The full-length RyR-1 cDNA was isolated from a rabbit skeletal muscle cDNA library as 9 overlapping clones (identified with probes from the published cDNA sequence²⁶) and was inserted, preceded by a Kozak initiation sequence (GCCGCC), into the XbaI site of pRM421 (gift of R. Mortensen, Brigham and Women's Hospital), to produce pRyR/Hygro, which confers hygromycin resistance and controls the insert with a phos-



phoglycerate kinase promoter. Preparation of normal ($+/\text{mdg}$) and dysgenic (mdg/mdg) myotubes was as described⁹. For dyspedic myotubes, dissociated cells from limb muscle were prepared²⁷ from newborn ($\text{ryr1}/\text{ryr1}$) mice. After proliferation in primary culture, cells were removed from culture dishes, and frozen at -80°C in Ham's F-10 with 20% fetal bovine serum (FBS) and 10% DMSO. Cells were thawed, cultured in DMEM with 20% FBS, and transferred after 5 d to DMEM with 5% heat-inactivated horse serum. Myotubes were injected⁵ with cDNA 6–8 d after initial plating, and examined 1–3 d later. Fluorescence changes (in arbitrary units) were measured after loading myotubes with Fluo-3 AM as previously described²⁸.

non-injected dyspedic myotubes were like that shown in Fig. 3b in exhibiting a small but measurable slow L-type current, whereas others completely lacked such current (Fig. 3e). Expression of RyR-1 in dyspedic myotubes caused a dramatic increase in the peak Ca^{2+} -current density (Fig. 3d), but the average ($6.2 \pm 2.5 \text{ pA/pF}$, $n = 12$) was still smaller than for normal myotubes, presumably because RyR-1 was only expressed in the vicinity of the one or two injected nuclei. The increased L-type current was attributable to the skeletal isoform of the DHPR, as injection of RyR-1 c-DNA did not alter L-type current density in dysgenic myotubes, which lack an intact gene for the skeletal DHPR, and because the current in RyR-1-expressing myotubes was blocked by $1 \mu\text{M}$ nifedipine (data not shown).

To investigate further the effect of RyR-1 on L-type Ca^{2+} -current density, we measured immobilization-resistant charge movements (Table 1). As previously described, the maximum immobilization-resistant charge movement (Q_{max}) was much smaller in dysgenic myotubes than in normal myotubes^{10,11}, consistent with the hypothesis that the DHPR produces a large fraction of the immobilization-resistant charge movement in normal myotubes. Values of Q_{max} were similar in dyspedic and RyR-1-expressing myotubes and close to the value for normal myotubes, suggesting that a significant density of DHPRs is present in the sarcolemma of dyspedic myotubes and that expression of RyR-1 does not alter this density. Rather, RyR-1 expression causes an increase in the ratio of L-type conductance to charge (Table 1), suggesting that DHPRs function poorly as Ca^{2+} -conducting channels in the absence of RyR-1.

Our results are consistent with a direct, reciprocal, energetic interaction between RyR-1 and the skeletal DHPR, an idea also supported by two earlier results. First, the open probability of skeletal DHPRs reconstituted into lipid bilayers has been reported to be increased roughly 2.5-fold by the addition of purified RyR-1 to the bathing solution¹². As well as affecting open probability, RyR-1 may stabilize higher-conductance states

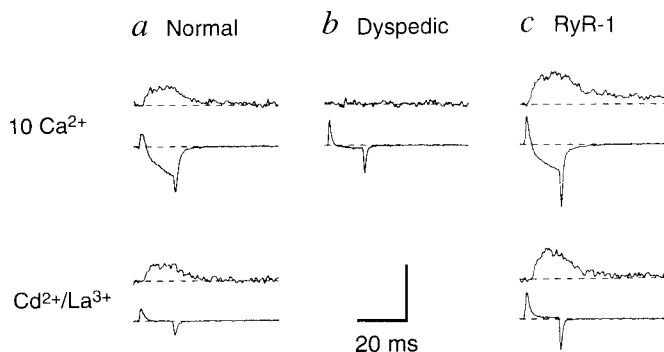


FIG. 2 Simultaneously recorded Ca^{2+} transients (upper trace of each pair) and membrane currents (lower trace of each pair) from normal (a), dyspedic (b) and RyR-1-expressing dyspedic myotubes (c). Data were obtained with an external solution that either permitted (10 mM Ca^{2+}) or blocked (0.5 mM Cd^{2+} and 0.3 mM La^{3+}) ionic current through Ca^{2+} channels. Expression of RyR-1 restored both Ca^{2+} current and depolarization-induced Ca^{2+} transients. a, Cell 598, linear capacitance (C) = 197 pF ; b, cell 444, $C = 395 \text{ pF}$; c, cell 590, $C = 690 \text{ pF}$. The test potential was $+30 \text{ mV}$ for 15 ms . Vertical calibration corresponds to 1.5 nA (a) or 3 nA (b, c) for currents, and $0.2 \Delta F/F$ for Ca^{2+} -induced fluorescence changes.

METHODS. For simultaneous measurement of membrane currents and Ca^{2+} transients, cells were held at -80 mV and stimulated with the prepulse protocol¹¹ as described previously¹⁰. The patch pipette contained (mM) 140 Cs-aspartate , 5 MgCl_2 , $10 \text{ Cs}_2\text{-EGTA}$ (20 nM free Ca^{2+}), 10 HEPES buffer ($\text{pH } 7.4$ with CsOH) and $5 \text{ Na}_2\text{-ATP}$ (for measurement of Ca^{2+} transients, 0.2 mM pentapotassium-Fluo-3 was also present). The bathing solution for measurement of Ca^{2+} currents was as described previously¹⁰; for measurements with ionic Ca^{2+} -channel current blocked, the bath contained (mM) a reduced Ca^{2+} concentration (2 rather than 10), 8 Mg^{2+} , 0.5 Cd^{2+} and 0.3 La^{3+} . Temperature was $20\text{--}22^\circ\text{C}$.

of the skeletal L-type channel, because channel activity associated with skeletal DHPRs in lipid bilayers (which obviously lack RyR-1) exhibits subconductance states^{13–15} not observed in intact muscle cells^{16–18}. The second result is that perchlorate both stabilizes the open state of RyR-1 and causes a hyperpolarizing shift in charge movement in skeletal muscle¹⁹ (where the DHPR and RyR-1 interact), but has little effect in systems (for example, cardiomyocytes) where the DHPR does not interact with RyR-1 (ref. 20). Unlike perchlorate, expression of RyR-1 in dyspedic myotubes did not significantly ($P > 0.01$) shift the voltage dependence of charge movement (Table 1).

The effect of RyR-1 on Ca^{2+} -channel activity may not involve a direct interaction with the DHPR $\alpha 1$ subunit. For example, RyR-1 might promote the association of $\alpha 1$ with accessory subunits (such as $\beta 1$) that could increase open probability. Alternatively, Ca^{2+} released from the sarcoplasmic reticulum might feed back during excitation–contraction coupling to potentiate the skeletal L-type current²¹.

Freeze–fracture studies demonstrate that junctional tetrads, which seem to be composed of DHPRs²², are absent in dyspedic muscle²³. Conceivably, the absence of RyR-1 could impair the biosynthesis or membrane insertion of DHPRs, but the small reduction of Q_{max} in dyspedic compared with normal myotubes (Table 1) indicates that this is not an important effect. Rather the Q_{max} values would seem to indicate that DHPRs are present but do not organize into tetrads in the absence of RyR-1 (which is a

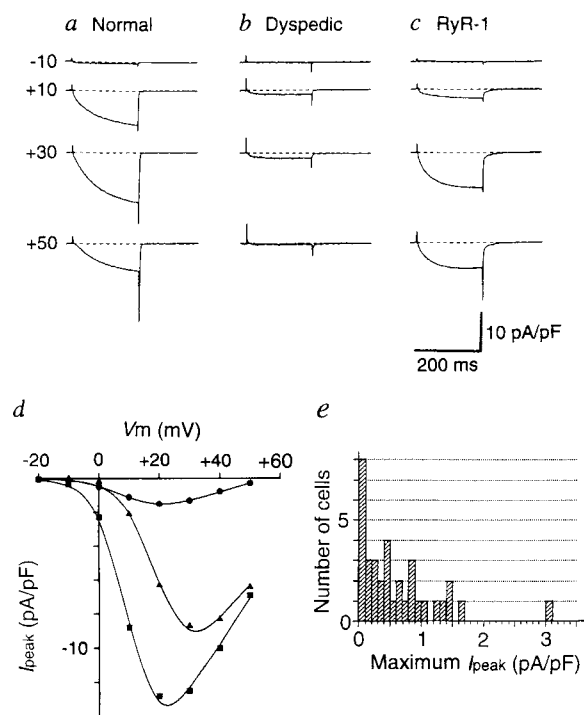


FIG. 3 Ca^{2+} currents recorded from normal (a), dyspedic (b) and RyR-1-expressing dyspedic myotubes (c). d, The peak current for these three cells, normalized by linear cell capacitance (I_{peak}), is plotted as a function of membrane voltage, V_m (normal $\blacksquare$, dyspedic $\bullet$, RyR-1-expressing $\blacktriangle$). e, Histogram of maximum I_{peak} in non-injected dyspedic cells. The measured current amplitudes in dyspedic cells may include contributions from a rapidly activating Ca^{2+} current (I_{dys}) that is also present in dysgenic muscle cells²⁹. Because I_{dys} is rapidly activating, its presence would reduce the apparent time constant of activation (τ_{act} ; ref. 30). τ_{act} values for the maximal current were $68.9 \pm 12.2 \text{ ms}$ ($n = 14$), $21.5 \pm 10.4 \text{ ms}$ ($n = 16$) and $47.9 \pm 10.9 \text{ ms}$ ($n = 12$) in normal, dyspedic and RyR-1-expressing myotubes, respectively. Values are given as mean $\pm$ s.d. One dyspedic myotube, which was found to possess a very rapidly activating Ca^{2+} current ($\tau_{\text{act}} < 7 \text{ ms}$), was excluded from further analysis. a, Cell 599, $C = 252 \text{ pF}$; b, cell 734, $C = 180 \text{ pF}$; c, cell 564, $C = 290 \text{ pF}$.

homotetrameric protein). Whatever the mechanism of RyR-1 expression on L-type current density, our results may explain why the skeletal L-type channel has been difficult to express in non-muscle systems, in which interactions between the DHPR and RyR-1 do not occur. □

Received 7 August 1995; accepted 5 January 1996.

1. Fleischer, S. & Inui, M. A. *Rev. Biophys. biophys. Chem.* **18**, 333–364 (1989).
2. McPherson, P. S. & Campbell, K. P. *J. biol. Chem.* **268**, 13765–13768 (1993).
3. Takeshima, H. et al. *Nature* **369**, 556–559 (1994).
4. Rios, E. & Brum, G. *Nature* **325**, 717–720 (1987).
5. Tanabe, T., Beam, K. G., Powell, J. A. & Numa, S. *Nature* **336**, 134–139 (1988).
6. Buck, R., Zimanyi, I., Abramson, J. J. & Pessah, I. N. *J. biol. Chem.* **267**, 23560–23567 (1992).
7. Takeshima, H. et al. *EMBO J.* **14**, 2999–3006 (1995).
8. Giannini, G., Conti, A., Mammarella, S., Scorbogna, M. & Sorrentino, V. J. *Cell Biol.* **128**, 893–904 (1995).
9. Penner, R., Neher, E., Takeshima, H., Nishimura, S. & Numa, S. *FEBS Lett.* **259**, 217–221 (1989).
10. Garcia, J., Tanabe, T. & Beam, K. G. *J. gen. Physiol.* **103**, 125–147 (1994).
11. Adams, B. A., Tanabe, T., Mikami, A., Numa, S. & Beam, K. G. *Nature* **346**, 569–572 (1990).
12. Ma, J., Hosey, M. M. & Rios, E. *Biophys. J.* **59**, 201a (1991).
13. Ma, J. & Coronado, R. *Biophys. J.* **53**, 387–395 (1988).
14. Yatani, A. et al. *J. biol. Chem.* **263**, 9887–9895 (1988).

15. Mejia-Alvarez, R., Fill, M. & Stefani, E. *J. gen. Physiol.* **97**, 393–412 (1991).
16. Caffrey, M. J., Brown, M. A. & Schneider, D. M. *Science* **236**, 570–573 (1987).
17. Winger, D. B. & Lansman, B. J. *J. Physiol., Lond.* **425**, 563–578 (1990).
18. Dirksen, R. T. & Beam, K. G. *J. gen. Physiol.* **105**, 227–247 (1995).
19. Gonzalez, A. & Rios, E. *J. gen. Physiol.* **102**, 373–421 (1993).
20. Ma, J. et al. *J. gen. Physiol.* **102**, 423–448 (1993).
21. Feldmeyer, D., Melzer, W., Pohl, B. & Zöllner, P. *Pflügers Arch. ges. Physiol.* **425**, 54–61 (1993).
22. Takekura, H., Bennett, L., Tanabe, T., Beam, K. G. & Franzini-Armstrong, C. *Biophys. J.* **67**, 793–803 (1994).
23. Takekura, H., Nishi, M., Noda, T., Takeshima, H. & Franzini-Armstrong, C. *Proc. natn. Acad. Sci. U.S.A.* **92**, 3381–3385 (1995).
24. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. *J. Embryol. exp. Morph.* **87**, 27–45 (1985).
25. Gossler, A., Doetschman, T., Korn, R., Serfling, E. & Kemler, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9065–9069 (1986).
26. Takeshima, H. et al. *Nature* **339**, 439–445 (1989).
27. Rando, T. A. & Blau, H. M. *J. Cell Biol.* **125**, 1275–1287 (1994).
28. Garcia, J. & Beam, K. G. *J. gen. Physiol.* **103**, 107–123 (1994).
29. Adams, B. A. & Beam, K. G. *J. gen. Physiol.* **94**, 429–444 (1989).
30. Tanabe, T., Adams, B. A., Numa, S. & Beam, K. G. *Nature* **352**, 800–803 (1991).

ACKNOWLEDGEMENTS. We thank N. Chaudhari for helpful discussion and R. Morris, S. Mukherjee and S. Schiefer for technical help. This research was supported by grants from MDA and the BWH Anesthesia Foundation (P.D.A.), the Boston Children's Heart Foundation (P.D.A. and H.T.N.) and the NIH (K.G.B.). R.T.D. was supported by an NRSA from the NIH.

Requirement for ceramide-initiated SAPK/JNK signalling in stress-induced apoptosis

Marcel Verheij[†]*, Ron Bose[‡], Xin Hua Lin[‡], Bei Yao[‡], W. David Jarvis[§], Steven Grant[§], Michael J. Birrer^{||}, Eva Szabo^{||}, Leonard I. Zon[¶], John M. Kyriakis[#], Adriana Haimovitz-Friedman^{*}, Zvi Fuks^{*} & Richard N. Kolesnick^{†*}

* Department of Radiation Oncology and ‡ Laboratory of Signal Transduction, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

§ Department of Medicine, Division of Hematology–Oncology, Medical College of Virginia, 1101 East Marshall Street, Richmond, Virginia 23298, USA

|| Biomarkers and Prevention Research Branch, Division of Cancer Prevention and Control, National Cancer Institute, 9610 Medical Center Drive, Rockville, Maryland 20850, USA

¶ Division of Hematology–Oncology, Children's Hospital and Dana-Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Howard Hughes Medical Institute, Children's Hospital, 300 Longwood Avenue, Enders 780, Boston, Massachusetts 02115, USA

Diabetes Research Laboratory, Medical Services, Massachusetts General Hospital East, 149 13th Street, Charlestown, Massachusetts 02129, USA

THE induction of programmed cell death, or apoptosis, involves activation of a signalling system, many elements of which remain unknown¹. The sphingomyelin pathway, initiated by hydrolysis of the phospholipid sphingomyelin in the cell membrane to generate the second messenger ceramide^{2,3}, is thought to mediate apoptosis in response to tumour-necrosis factor (TNF)- α ^{2,3}, to Fas ligand⁴ and to X-rays⁵. It is not known whether it plays a role in the stimulation of other forms of stress-induced apoptosis. Given that environmental stresses also stimulate a stress-activated protein kinase (SAPK/JNK)^{6–12}, the sphingomyelin and SAPK/JNK signalling systems may be coordinated in induction of apoptosis. Here we report that ceramide initiates apoptosis through the SAPK cascade and provide evidence for a signalling mechanism that integrates cytokine- and stress-activated apoptosis.

We studied stress effects in two cell types, U937 human monocytic leukaemia and bovine aortic endothelial (BAE) cells. Ceramide-initiated apoptosis has been investigated in these systems^{5,13}. Exposure of U937 cells to ionizing radiation (10 Gy), hydrogen peroxide (1 mM), ultraviolet-C radiation (10,000 $\mu\text{J cm}^{-2}$), heat shock (45 °C) and TNF- α (10 nM) rapidly induces ($P < 0.05$) an increase in ceramide above a basal level of 80 pmol per 10^6 cells (Fig. 1a), with a concomitant quantitative reduction in sphingomyelin (data not shown). A similar result was obtained with BAE cells (not shown). The effects are dose-dependent (Fig. 1b).

Exposure to stress or C2-ceramide (100 μM) induced the morphological and biochemical features typical of apoptosis¹, measured by using the DNA-specific fluorochrome bisbenzimidazole¹⁴ and confirmed by TUNEL assay¹⁵ (results not shown). Apoptotic changes were detected in U937 cells (Fig. 2a) and BAE cells (not shown) after 6 hours of stimulation and were time- and dose-dependent. Apoptosis was specific for ceramide as C2-dihydroceramide which lacks the *trans* double bond at C4–5 of the sphingoid base backbone, was ineffective. LD₅₀ (the lethal dose for 50% of cells) values for stress-induced apoptosis were closely correlated with ED₅₀ (half the maximum effective dose) values for ceramide generation in both cell types (Fig. 1b).

As environmental stress may activate the SAPK cascade^{6–8,10–12,16}, we investigated the effect of ceramide on this system. The SAPK/JNK pathway involves sequential activation of the proteins MEKK1, SEK1, SAPK and c-Jun^{6,9–12}. We stimulated U937 cells with C2-ceramide or *Staphylococcus aureus* sphingomyelinase, immunoprecipitated SAPK and used it to phosphorylate an N-terminal c-Jun construct (GST–c-Jun-1-135)⁶. Figure 3a shows that C2-ceramide and sphingomyelinase induced concentration-dependent SAPK activation to as much as 40-fold of control; similar results were obtained in BAE cells (not shown). Environmental stress also activated SAPK to 20- to 40-fold of control in U937 and BAE cells (Figs 2c, 4c). In contrast, equimolar concentrations (100 μM) of other lipid second messengers, 1,2-dioctanoyl-*sn*-glycerol, 1,2-dioctanoyl-*sn*-glycero-3-phosphate, arachidonic acid and C2-dihydroceramide, which do not initiate apoptosis in these cells¹³, failed to activate SAPK (Fig. 3b). Thus ceramide, like stress, activates SAPK.

The role of SAPK in stress-induced apoptosis was assessed with TAM-67 (ref. 17), a dominant-negative c-Jun mutant lacking the N terminus. Figure 2a shows that stress- and C2-ceramide (100 μM)-induced apoptosis was inhibited in U937/TAM-67 cells; results were similar from U937/TAM-67 clones from three independent stable transfections. Transient TAM-67 expression also inhibited stress- and ceramide-induced apoptosis in U937 cells and BAE cells by 40 and 56% at 16 h respectively. Ceramide generation (Fig. 2b) and SAPK activation (Fig. 2c) were not compromised in

[†] Present address: Department of Radiation Therapy, Antoni v. Leeuwenhoekhuis/The Netherlands Cancer Institute, Amsterdam, The Netherlands.

^{*} To whom correspondence should be addressed.

stressed cells stably expressing TAM-67. Hence, inhibition of the apoptotic response to environmental stress occurs at a level downstream of SAPK. Induction of apoptosis by calphostin C and cytosine arabinoside were not inhibited in U937/TAM-67 cells (data not shown), indicating that the systems that effect apoptosis are intact in these cells.

We also transiently transfected U937 cells with a dominant-negative kinase-inactive SEK1 construct, pEBG-SEK1(K → R)¹⁰. Transfection efficiency was $39 \pm 5\%$ ($n = 9$). In cells overexpressing SEK1(K → R), activation of endogenous SAPK in response to C2-ceramide (100 μM) or H_2O_2 (1 mM) was inhibited by 50% (Fig. 4a, top), whereas activation of co-transfected HA-tagged SAPK was blocked (Fig. 4a, bottom), indicating that signalling was specifically inhibited in the transfected cells. Apoptosis in response to C2-ceramide was accordingly reduced at 8 and 16 h (Fig. 4b): in four experiments, reduction was 53% at 8 h and 58% at 16 h ($P < 0.005$ versus wild type plus C2-ceramide), an effect that persisted for 36 h. Transient SEK1(K → R) transfection also reduced apoptosis after exposure to ionizing radiation (10 Gy), H_2O_2 (1 mM), ultraviolet-C radiation (10,000 $\mu\text{J cm}^{-2}$), heat shock (45 °C) and TNF- α (10 nM) by a mean of 55% at 8 and 16 h. In BAE cells, transient SEK1(K → R) overexpression gave 43% inhibition of stress-induced apoptosis at 16 h. Together these results indicate that expression of stress-induced apoptosis in U937 and BAE cells is signalled by ceramide and requires a functional SAPK/JNK cascade. Consistent with these observations, it has recently been reported that, in PC12 cells deprived of nerve-growth factor (NGF), apoptosis is mediated via the SAPK/JNK system¹⁸.

Our results shed light on the dual function of TNF- α in

mammalian systems. TNF- α appears to induce two major effects through ceramide, namely inflammation and apoptosis⁵. Although we have shown that ceramide signals apoptosis by means of the SAPK cascade in U937 and BAE cells, in HL-60 cells ceramide links the 55K TNF- α receptor to the inflammatory response through the MAPK cascade¹⁹. In BAE cells, C2-ceramide and stress do not significantly activate MAPK (except minimally by heat) (Fig. 4c). In contrast, the phorbol ester TPA (200 ng ml^{-1}) or basic fibroblast growth factor (bFGF; 1 ng ml^{-1}), both mitogens for BAE cells⁵, increased MAPK activation 50-fold. Similarly, stress did not activate the MAPK cascade in U937 cells. Further, transient overexpression of the dominant-negative kinase-inactive Raf-1 mutant RSV-Raf-C4 (ref. 20), or the catalytically inactive dominant-negative MEK1 mutant CMV-MKK-K97M²¹ in U937 cells (Fig. 4d) or BAE cells (not shown) did not affect ceramide-induced apoptosis. The same results were obtained using all five types of stress (not shown). Adequate expression of these inhibitory constructs was demonstrated by complete suppression of MAPK activation following stimulation of BAE cells with TPA (200 ng ml^{-1}) and bFGF (1 ng ml^{-1}) (Fig. 4e). These results indicate that stress-induced apoptosis in U937 and BAE cells is signalled through SAPK and does not involve the MAPK pathway. Whether ceramide signalling through the MAPK and SAPK cascades is cell-type-specific, or whether a single cell can shift between these two systems, is not known. However, distinct TNF receptor domains are known to link acid and neutral sphingomyelinases to different downstream signalling pathways²², with the neutral sphingomyelinase specifically coupled to the MAPK cascade.

The mechanism by which c-Jun mediates apoptosis is unknown. Apoptosis of rat neuronal cells after NGF deprivation is inhibited

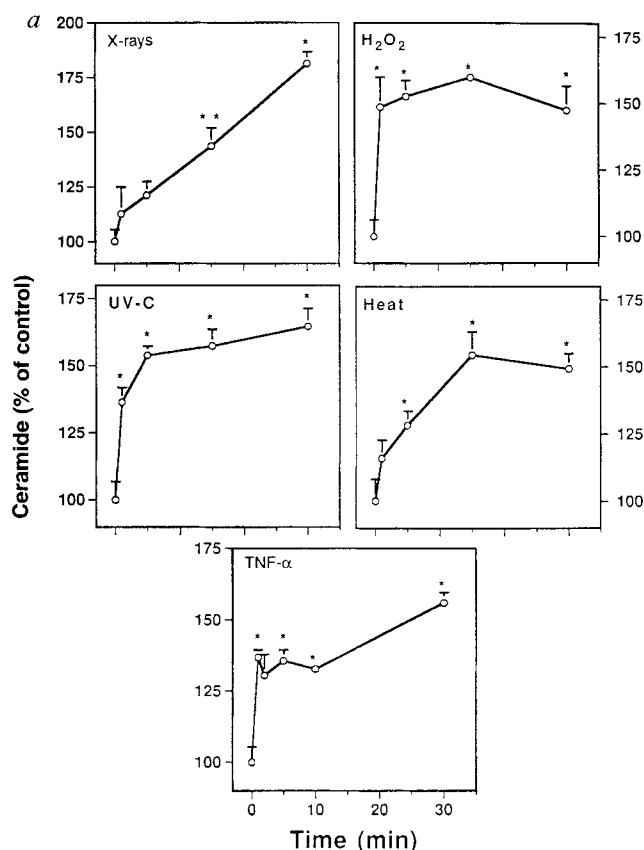


FIG. 1 Stress-induced ceramide generation and apoptosis in U937 cells. a, Time course of ceramide generation in response to X-rays (10 Gy), H_2O_2 (1 mM), ultraviolet-C radiation (UV-C; 10,000 $\mu\text{J cm}^{-2}$), heat shock (45 °C) and TNF- α (10 nM). b, ED₅₀ for stress-induced ceramide generation at 30 min and LD₅₀ for apoptosis at 12 h in U937 and BAE cells.

b

	U937 cells		BAE cells	
	ED ₅₀	LD ₅₀	ED ₅₀	LD ₅₀
Ionizing radiation (Gy)	5	3	1.5	3
Hydrogen peroxide (mM)	0.6	0.4	0.6	0.7
Ultraviolet-C radiation ($\mu\text{J cm}^{-2}$)	10,000	8,000	5,000	3,000
Heat shock (°C)	42	41	43	43
TNF- α (nM)	0.5	1.0	0.9	1.5

METHODS. a, U937 cells ($0.6 \times 10^6 \text{ ml}^{-1}$) and confluent monolayers of BAE cells, cultured as described^{5,13}, were incubated in serum-free medium and stressed after 20 h. Samples were irradiated on ice in a ¹³⁷Cs gamma-cell-40 chamber (Atomic Energy of Canada) at a dose rate of 100 cGy min⁻¹. Hydrogen peroxide was diluted in water. UV-C (at 3,000 $\mu\text{W cm}^{-2}$, wavelength 254 nm) was delivered using a Stratalinker UV crosslinker. Heat shock was in a hot-water bath with regulated air and CO₂ flow. TNF- α (R&D Systems) was diluted in 0.1% BSA. Diluent was without effect. Ceramide was extracted and quantified using *Escherichia coli* DAG kinase (Calbiochem) as described⁵. Data (mean $\pm$ s.d.) were derived from three separate determinations from two experiments in which * $P < 0.01$ and ** $P < 0.05$ compared with controls (Student's *t*-test). b ED₅₀ for ceramide generation and LD₅₀ for apoptosis were determined using the following dose ranges: ionizing radiation, 0.5–10 Gy; H_2O_2 , 0.1–10 mM; UV-C, 1,000–100,000 $\mu\text{J cm}^{-2}$; heat, 41–46 °C; and TNF- α , 0.1–30 nM. Apoptosis was quantified by staining with bisbenzimide (Hoechst-33258, Sigma) as described¹⁴. 500 cells were scored for each data point. Data are expressed as mean values from duplicate determinations from 2 experiments.

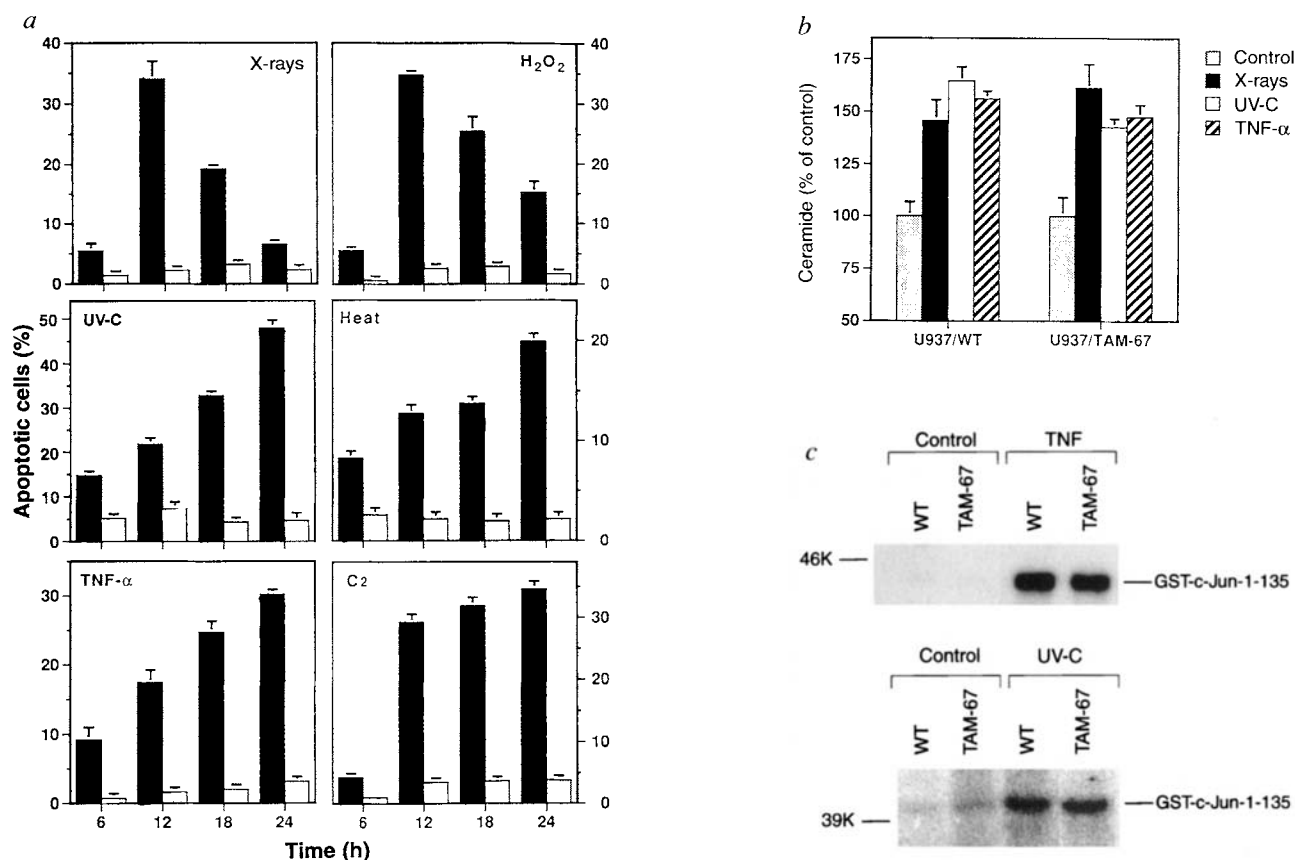
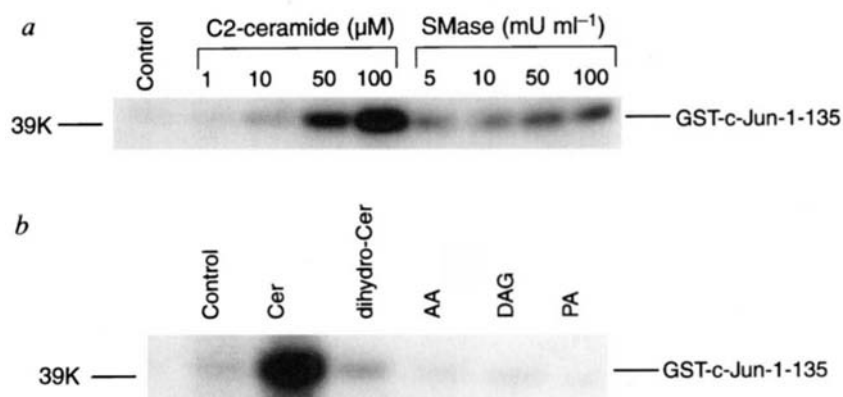


FIG. 2 A dominant-negative N-terminal deletion mutant of c-Jun (TAM-67) blocks stress-induced apoptosis but does not inhibit generation of ceramide or SAPK activation in U937 cells. **a**, Time course of apoptotic nuclear changes in response to X-rays (5 Gy), H_2O_2 (1 mM), ultraviolet radiation (10,000 $\mu J cm^{-2}$), heat shock (45 °C for 30 min), TNF- α (10 nM) and C2-ceramide (100 μM) in U937/WT ('wild-type' strain; black bars) and U937/TAM-67 (white bars). **b**, Ceramide levels in U937/WT and U937/TAM-67 at 30 min after X-ray radiation (10 Gy), ultraviolet radiation (10,000 $\mu J cm^{-2}$) or TNF- α (10 nM). **c**, Activation of SAPK at 20 min by TNF- α (1 nM; top panel) or UV-C (10,000 $\mu J cm^{-2}$; bottom panel) in U937/WT and U937/TAM-67. Relative molecular mass (M_r) markers are indicated on the left.

METHODS. Stable transfection of TAM-67 (ref. 17) into U937 cells was achieved using the pMexMth-TAM-67 construct, which expresses the bacterial neomycin-resistance gene (*neo*^r), as described²⁹. Transfectants were maintained in medium supplemented with the neomycin analogue geneticin (G418, 400 $\mu g ml^{-1}$; GIBCO). There was no difference in signalling of apoptosis in wild-type U937 cells and cells transfected with empty vector, so these lines were used interchangeably. Detection of apoptotic nuclear changes by bisbenzamide staining (**a**), measurements of cellular ceramide (**b**) and identification of SAPK activation (**c**) after cellular stress were as described for Figs 1 and 3. Data in **a** (mean $\pm$ range) are derived from duplicate determinations from 2 experiments; in **b**, (mean $\pm$ s.d.) are derived from 3 separate determinations from 2 experiments.

FIG. 3 Ceramide induces SAPK activation in U937 cells. Autoradiograph after SDS-PAGE showing the effect of C2-ceramide and sphingomyelinase (**a**) and of other lipid second messengers (**b**) on SAPK activation. The phosphorylated GST-c-Jun-1-135 band is indicated.

METHODS. U937 cells were handled as for Fig. 1, followed by 20 min treatment with C2-ceramide (1–100 μM), *S. aureus* neutral sphingomyelinase (SMase, 5–100 $mU ml^{-1}$, Sigma), C2-dihydroceramide (dihydro-Cer, 100 μM), arachidonic acid (AA, 100 μM), 1,2-dioctanoyl-*sn*-glycerol (DAG, 100 μM) or 1,2-dioctanoyl-*sn*-glycerol-3-phosphate (PA, 100 μM) (Biomol). Cells were then lysed in lysis buffer (20 mM HEPES, pH 7.4, 2 mM EGTA, 50 mM β -glycerophosphate, 1 mM DTT, 1 mM Na_3VO_4 , 1% Triton X-100, 10% glycerol, 2 μM leupeptin, 10 $\mu g ml^{-1}$ soy bean trypsin inhibitor, 400 μM PMSF), solubilized on ice and the nuclei-free supernatant was normalized for protein content and immunoprecipitated with anti-SAPK antibody-conjugated Sepharose beads⁶. To 41.5 μl SAPK-bound beads in assay buffer were added 4 μl 1 $mg ml^{-1}$



GST-c-Jun-1-135 (ref. 6), and 4.5 μl Mg-ATP mix (20 mM $MgCl_2$, 25 μM ATP and 1 μl [γ -³²P]ATP (3,000 Ci mmol⁻¹). The reaction was terminated after 20 min at 30 °C by addition of Laemmli buffer and the products resolved by 10% SDS-PAGE and autoradiography.

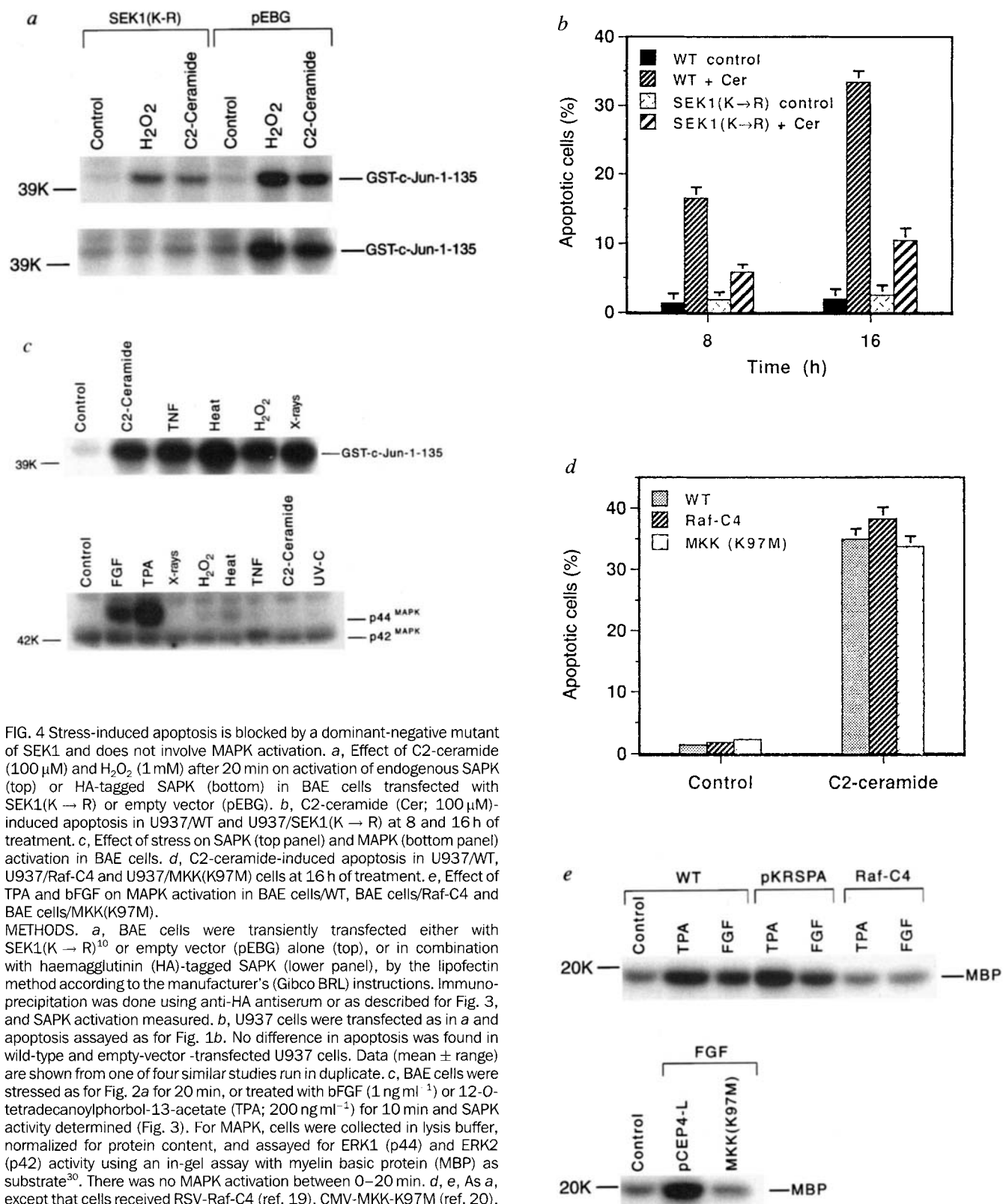


FIG. 4 Stress-induced apoptosis is blocked by a dominant-negative mutant of SEK1 and does not involve MAPK activation. **a**, Effect of C2-ceramide (100 μ M) and H₂O₂ (1 mM) after 20 min on activation of endogenous SAPK (top) or HA-tagged SAPK (bottom) in BAE cells transfected with SEK1(K $\rightarrow$ R) or empty vector (pEBG). **b**, C2-ceramide (Cer; 100 μ M)-induced apoptosis in U937/WT and U937/SEK1(K $\rightarrow$ R) at 8 and 16 h of treatment. **c**, Effect of stress on SAPK (top panel) and MAPK (bottom panel) activation in BAE cells. **d**, C2-ceramide-induced apoptosis in U937/WT, U937/Raf-C4 and U937/MKK(K97M) cells at 16 h of treatment. **e**, Effect of TPA and bFGF on MAPK activation in BAE cells/WT, BAE cells/Raf-C4 and BAE cells/MKK(K97M).

METHODS. **a**, BAE cells were transiently transfected either with SEK1(K $\rightarrow$ R)¹⁰ or empty vector (pEBG) alone (top), or in combination with haemagglutinin (HA)-tagged SAPK (lower panel), by the lipofectin method according to the manufacturer's (Gibco BRL) instructions. Immunoprecipitation was done using anti-HA antiserum or as described for Fig. 3, and SAPK activation measured. **b**, U937 cells were transfected as in **a** and apoptosis assayed as for Fig. 1b. No difference in apoptosis was found in wild-type and empty-vector-transfected U937 cells. Data (mean $\pm$ range) are shown from one of four similar studies run in duplicate. **c**, BAE cells were stressed as for Fig. 2a for 20 min, or treated with bFGF (1 ng ml⁻¹) or 12-O-tetradecanoylphorbol-13-acetate (TPA; 200 ng ml⁻¹) for 10 min and SAPK activity determined (Fig. 3). For MAPK, cells were collected in lysis buffer, normalized for protein content, and assayed for ERK1 (p44) and ERK2 (p42) activity using an in-gel assay with myelin basic protein (MBP) as substrate³⁰. There was no MAPK activation between 0–20 min. **d**, **e**, As **a**, except that cells received RSV-Raf-C4 (ref. 19), CMV-MKK-K97M (ref. 20), or their respective empty vectors (pKRSPA or pCEP4-L). Data (mean $\pm$ range) in **d** are from duplicate determinations from 2 independent experiments. For **e**, BAE cells were cotransfected with HA-tagged ERK1 and treated with diluent (control), TPA (200 ng ml⁻¹) or bFGF (1 ng ml⁻¹) for 10 min. After immunoprecipitation with anti-HA serum, MAPK activity was determined by an immune complex kinase assay (Fig. 3) using MBP as substrate. Transfection efficiency was monitored with a mammalian green fluorescent protein (GFP) expression construct (pGFP-N1, Clontech). In both cell lines, maximal expression to 40–45% of the population occurred at 48 h after transfection and persisted for 72 h. Stimulation experiments were initiated at 38 h post-transfection.

by a c-Jun dominant-negative mutant²³. As apoptosis is also inhibited by actinomycin D and cycloheximide²⁴ it was suggested that c-Jun might regulate genes as yet unknown, that are involved in cell proliferation, and that strong growth-promoting signals in non-cycling neurons might trigger apoptosis²⁵. However, chemical inhibition of transcription and protein synthesis in U937 cells does not inhibit apoptosis²⁵. Further, there was no correlation between apoptosis and cycling of U937 and BAE cells. The BAE cells used here are non-cycling⁵, whereas the U937 cells are cycling. An alternative hypothesis might stem from the capacity of c-Jun to complex with different proteins, such as ATF-2, NF-IL6 and NFAT²⁶. Perhaps, activated c-Jun may sequester and regulate unknown proteins required in the apoptotic response.

Although we have provided evidence for the coordinated regulation of apoptosis via the sphingomyelin and SAPK pathways, this does not mean that this is the only system for mediating apoptosis. As U937/TAM-67 cells undergo cytosine arabinoside- and calphostin C-induced apoptosis, other mechanisms probably exist. SAPK activation need not always lead to apoptosis: in fact, like ceramide, activation of SAPK has been associated with stimulation of proliferative and differentiated functions^{7,27}. One suggestion is that the context of the ceramide signal determines the ultimate biological response²⁸. In this regard, ceramide-mediated apoptosis appears to be subject to transmodulatory control through 1,2-diacylglycerol (DAG)/protein kinase C; apoptosis initiated by TNF- α and ceramide analogues in U937 cells is inhibited by an increase in DAG concentration^{13,28}. Also, in Jurkat T cells, DAG converts an apoptotic response to ceramide into a proliferative signal²⁸. Thus, even within the same cell, an increase in ceramide may lead to different end points depending on the microenvironment in which the signal is generated. $\square$

Received 14 September 1995; accepted 18 January 1996.

1. Kerr, J. F. R. & Harmon, B. V. in *Apoptosis: The Molecular Basis of Cell Death* (eds Tormei, L. D. & Cope, F. O.) 5–29 (Cold Spring Harbor Laboratory Press, New York, 1991).
2. Pushkareva, M., Obeid, L. M. & Hannun, Y. A. *Immun. Today* **16**, 294–297 (1995).
3. Kolesnick, R. & Golde, D. W. *Cell* **77**, 325–328 (1994).
4. Gulbins, E. et al. *Immunity* **2**, 341–351 (1995).
5. Haimovitz-Friedman, A. et al. *J. exp. Med.* **180**, 525–535 (1994).
6. Kyriakis, J. M. et al. *Nature* **369**, 156–160 (1994).
7. Dérjard, B. et al. *Cell* **76**, 1025–1037 (1994).
8. Sluss, H. K., Barrett, T., Dérjard, B. & Davis, R. J. *Molec. cell. Biol.* **14**, 8376–8384 (1994).
9. Dérjard, B. et al. *Science* **267**, 682–685 (1995).
10. Sánchez, I. et al. *Nature* **372**, 794–798 (1994).
11. Yan, M. et al. *Nature* **372**, 798–800 (1994).
12. Lin, A. et al. *Science* **268**, 286–290 (1995).
13. Jarvis, W. D. et al. *J. biol. Chem.* **269**, 31685–31692 (1994).
14. Oberhammer, F. A. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5408–5412 (1992).
15. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. *J. Cell Biol.* **119**, 493–501 (1992).
16. Westwick, J. K., Bielawska, A. E., Dbaibo, G., Hannun, Y. A. & Brenner, D. A. *J. biol. Chem.* **270**, 22689–22692 (1995).
17. Brown, P. H., Chen, T. K. & Birrer, M. J. *Oncogene* **9**, 791–799 (1993).
18. Xia, Z., Dickens, M., Raingeaud, J., Davis, R. & Greenberg, M. E. *Science* **270**, 1326–1331 (1995).
19. Yao, B. et al. *Nature* **378**, 307–310 (1995).
20. Bruder, J. T., Heidecker, G. & Rapp, U. R. *Genes Dev.* **6**, 545–556 (1992).
21. Mansour, S. J. et al. *Science* **265**, 966–970 (1994).
22. Wiegmann, K., Schutze, S., Machleidt, T., Witte, D. & Krönke, M. *Cell* **78**, 1005–1015 (1994).
23. Ham, J. et al. *Cell* **14**, 927–939 (1995).
24. Johnson, E. M. & Deckwarth, T. L. A. *Rev. Neurosci.* **16**, 31–46 (1993).
25. Martin, S. J. *Trends Cell Biol.* **3**, 141–144 (1993).
26. Karin, M. *J. biol. Chem.* **270**, 16483–16486 (1995).
27. Su, B. et al. *Cell* **77**, 727–736 (1994).
28. Kolesnick, R. & Fuks, Z. *J. exp. Med.* **181**, 1949–1952 (1995).
29. Alani, R. et al. *Molec. cell. Biol.* **11**, 6286–6295 (1991).
30. Kameshita, I. & Fujisawa, H. *Analyt. Biochem.* **183**, 139–143 (1989).

ACKNOWLEDGEMENTS. We thank N. Ahn for SEK1(K → R) and R. S. Persaud for technical assistance. M.V. held a radiobiological fellowship from the Dutch Cancer Society (Koninkrijk Wilhelmina Fonds). This work was supported by an MSTP training grant and the Lee Friedman Memorial Fellowship (R.B.); the Charles A. Dana Foundation and the Michael and Ethel Cohen Fellowship Fund (X.H.L.); a National Research Service award (W.D.J.); the intramural research program of the NCI (M.J.B. and E.S.); and grants from the NCI (W.D.J. and S.G.), the NIH (L.L., J.M.K., A.H.-F., Z.F. and R.N.K.) and the US Army (breast cancer research grant to J.M.K.).

Telomerase activation by the E6 gene product of human papillomavirus type 16

Aloysius J. Klingelhutz*, Scott A. Foster* & James K. McDougall*†

* Program in Cancer Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA

† Department of Pathology, University of Washington, Seattle, Washington 98195, USA

ACTIVATION of telomerase, a ribonucleoprotein complex that synthesizes telomere repeat sequences^{1–4}, is linked to cell immortalization and is characteristic of most cell lines and tumours^{4,5–8}. Here we show that expression of the human papillomavirus type 16 (HPV-16) E6 protein activates telomerase in early-passage human keratinocytes and mammary epithelial cells. This activation was observed in cells pre-crisis, that is, before they became immortal, and occurred within one passage of retroviral infection with vectors expressing HPV-16 E6. Studies using HPV-16 E6 mutants showed that there was no correlation between the ability of the mutants to activate telomerase and their ability to target p53 for degradation, suggesting that telomerase activation by HPV-16 E6 is p53 independent. Keratinocytes expressing wild-type HPV-16 E6 have an extended lifespan, but do not become immortal⁹, indicating that telomerase activation and E6-mediated degradation of p53 are insufficient for their immortalization. These results show that telomerase activation is an intrinsic, but insufficient, component of transformation by HPV.

Telomeres, composed of TTAGGG repeats, have been shown to shorten as a function of age *in vivo*^{10–12} and during the passage of normal cells *in vitro*^{6,9,13,14}. This loss of telomere sequences is thought to contribute to senescence^{14,15}. Activation of a mechanism to restore telomeres, namely, telomerase activation, has been proposed to be critical event in the immortalization of human cells^{4–7}. We examined pre- and post-crisis human cervical keratinocytes (HCK) that expressed HPV-16 E6 and E7 together or separately from retroviral constructs^{16,17} with a view to establishing a temporal relationship between viral gene expression, immortalization and activation of telomerase. Telomerase activity was assayed by using a modified version of the telomere repeat amplification protocol (TRAP) assay⁸. Only those cells which expressed both E6 and E7 became immortal and exhibited high levels of telomerase activity, but we also detected telomerase activity in pre-crisis cells (Fig. 1a). This activity was limited to cells expressing E6, either alone or with E7. To determine how soon telomerase activation occurs after introduction of E6, we examined early-passage, pre-crisis human foreskin keratinocytes (HFK) within one passage of retroviral infection (Fig. 1b). Again, telomerase activity was detected and limited to those cells which expressed E6, indicating that telomerase activation occurred soon after the introduction of E6, and was not due to a rare genetic event in the cell population. E6-associated telomerase activation was also detected in early-passage pre-crisis human mammary epithelial cells (HMEC) (Fig. 1b). However, we could not detect telomerase activity in E6-expressing early-passage human fibroblasts (data not shown), which suggests that the E6-associated telomerase activation is cell-type specific.

The finding that telomerase activation is a feature of pre-crisis E6-expressing epithelial cells might be interpreted as contradicting a previous observation that such cells exhibit telomere shortening with passage in culture⁹. Serial dilution of cell extracts

from E6-expressing cells and subsequent TRAP assay analysis indicated that the amount of telomerase activity in pre-crisis cells was approximately 1/100 of the amount in post-crisis E6/E7-expressing immortal cells (Fig. 2). Cloning of the pre-crisis E6-transfected cervical cells resulted in the isolation of three clones in which telomerase activity could not be detected (data not shown). This indicates that measurable telomerase activity occurred in a subpopulation of the initial cultures, which would account for telomere shortening in the population overall.

To determine whether telomerase activation was p53 dependent, we used an array of HPV-16 E6 mutants that have been shown to exhibit differential ability to target p53 for degradation

when transfected into primary human cells¹⁸. Two mutants, the carboxy-terminal mutant HPV-16 E6 Δ 146–151 and the amino-terminal mutant HPV-16 E6 8S/9A/10T, both activated telomerase (Fig. 3a), but HPV-16 E6 Δ 146–151 targets p53 for degradation, whereas HPV16 E6 8S/9A/10T does not (Fig. 3a). Furthermore, the mutant HPV-16 E6 Δ 118–122 efficiently degrades p53 but does not activate telomerase. We also observed, in two of three independent experiments, low telomerase activity in cells expressing the low-risk HPV-6 E6 (Fig. 3b), which fails to target p53 for degradation (Fig. 3a and ref. 18). Thus telomerase activation appears to be independent of E6-mediated degradation of p53, and may result from another, as yet unknown, function of

FIG. 1 a, TRAP assay for telomerase activity in pre- and post-crisis human cervical keratinocytes (HCK) expressing HPV oncoproteins. Primary cells at passage (P) 3 were infected with LXS-based retroviral vectors expressing HPV-16 E6 alone, HPV-16 E7 alone, HPV-16 E6/E7 together, or control^{16,17}, selected with G418, and shown to express the appropriate HPV genes as described previously⁹. Normal HCK and LXS HCK senesced at P11, whereas HPV-16 E6 HCK and HPV-16 E7 HCK had an extended lifespan to P19 and P15, respectively, but then senesced⁹. The HPV-16 E6/E7 cells went into crisis at P19 to P20, and thereafter became immortal (that is, greater than 80 passages). The TRAP assay was performed as described^{8,21}, using 5 μ g of protein in a 50 μ l reaction. Telomerase synthesizes telomere repeats onto a non-telomeric ³²P end-labelled repeat TS (5'-AATCCGTCGAGCAGAGTT-3'), and these products are specifically amplified by polymerase chain reaction (PCR) with the downstream primer CX (5'-(CCCTTA)₃CCCTAA-3') and upstream labelled primer TS. A portion of the reaction (6 μ l) was run on a 7.5% non-denaturing acrylamide gel, which was dried and exposed to film for 2 days. The ladder observed in telomerase-positive samples represents 6-base-pair repeats with the smallest product representing 40 bp. The HPV-16 E6/E7-expressing cells that survived crisis (P27, P48 and P84) showed significant telomerase activity. Telomerase activity was also observed in pre-crisis cells that expressed E6 alone or in combination with E7 soon after transfection and selection (P8), but it was not observed in normal HCK, vector control or cells that expressed E7 alone. As reported previously¹⁻⁴, telomerase activity was shown to be sensitive to RNase by the addition of 0.5 μ g of RNase A in the elongation step ('-', without RNase; '+', with RNase). The (-) control lane represents a reaction performed with all the components except cell extract. Photographs were generated by scanning the autoradiograph using Sharp JX-325 scanner and Adobe photoshop software, labelling using Canvas 3.0 software, and printing on a Tektronix Phaser II S2X dye sublimation printer. b, TRAP assay for telomerase activity in newly infected human foreskin keratinocytes (HFK) and human mammary epithelial cells (HMEC). HFK cell lysates were collected for the telomerase assay after selection in G418 within one passage of retroviral infection; HMEC cells were collected within two passages. Immortalized HPV-16 E6/E7 HCK (P84) were used as a positive control. Cells transfected with HPV-16 E6, either alone or in combination with HPV-16 E7, expressed telomerase activity very early after transfection.

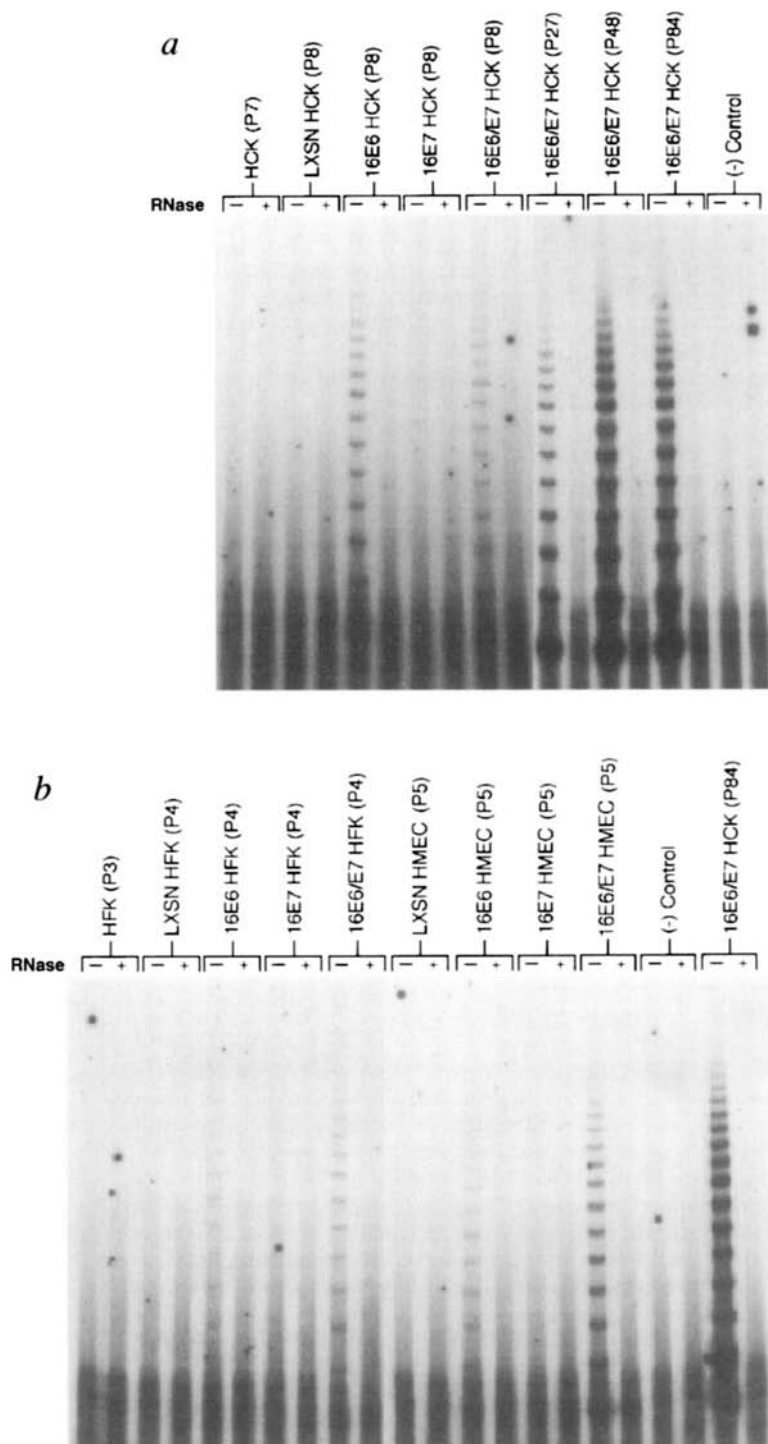


FIG. 2 Dilution analysis of telomerase-expressing cells to determine relative activity. TRAP assays were performed as in Fig. 1. Cell lysates from post-crisis HPV-16 E6/E7 HCK (P84), pre-crisis HPV-16 E6 HCK (P4), and pre-crisis HPV-16 E6/E7 HCK (P4) were diluted 1:10, 1:100, and 1:1,000 in lysis buffer, and the amounts shown were then subjected to the TRAP assay. A portion of each reaction (6 μ l) was then run on a 7.5% non-denaturing polyacrylamide gel, which was dried and exposed to film for 1 day. Visible telomerase activity was observed in the 1:1,000 dilution (0.005 μ g protein per reaction) in the post crisis immortalized cells but was not observed in the 1:1,000 or 1:100 (0.05 μ g per reaction) dilutions of the pre-crisis HPV-16 E6 and HPV-16 E6/E7 lysates, indicating that telomerase activity in the pre-crisis cells was approximately 1% of that in the post-crisis cells.

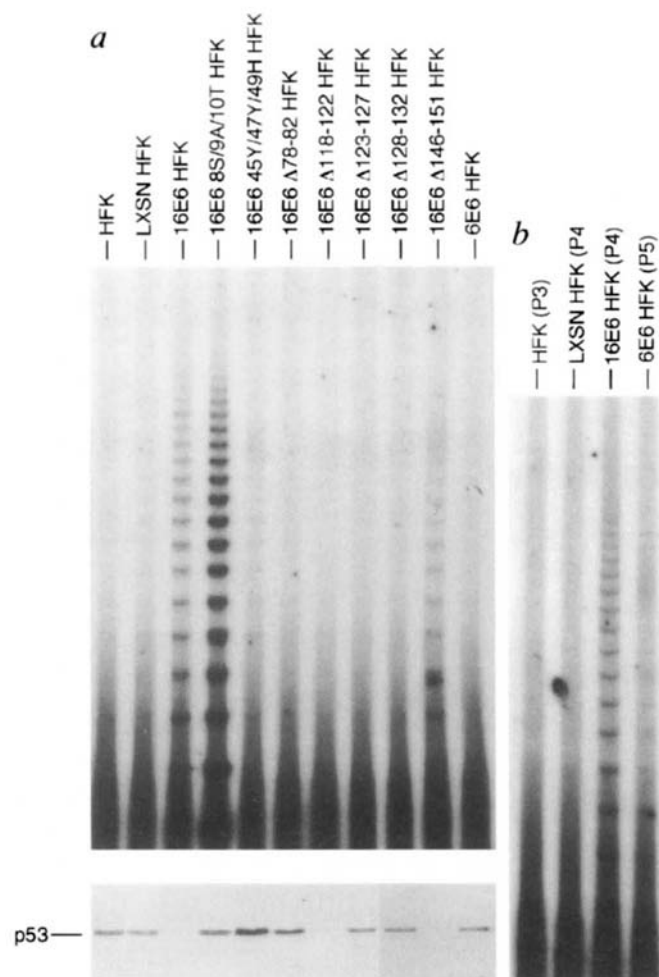
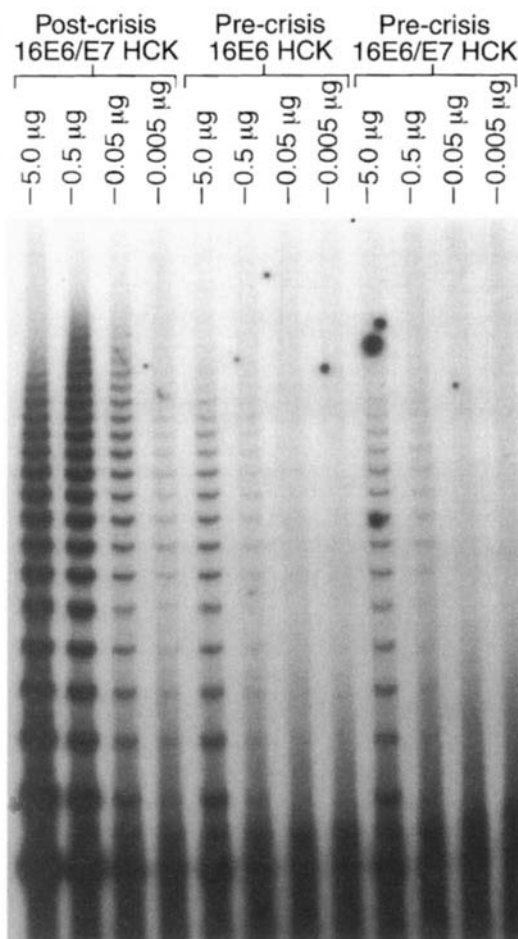


FIG. 3 *a*, TRAP analysis of HPV-16 E6 mutant-expressing cells. The E6 mutants and retroviral constructs have been described previously¹⁸. Primary early-passage HFK were infected with retroviral constructs as in Fig. 1 and examined for telomerase activity within one passage of infection. Samples were run as in Fig. 1, and the gel was exposed to film for 2 days. Western analysis (bottom) for p53 protein was performed as described previously¹⁸ using cell lysates collected within two passages of retroviral infection, and 20 μ g protein lysate was loaded in each lane. Photographs were generated as in Fig. 1. *b*, Trap analysis of HPV-6 E6-expressing cells. Primary early-passage HFK were infected with retroviral constructs as in Fig. 1 and examined for telomerase activity within two passages of infection.

E6. All of the E6 mutations that did not activate telomerase map to the interior portion of the protein in the zinc-finger domains and central region, suggesting that these regions of E6 are necessary for telomerase activation. Furthermore, the HPV-16 E6 8S/9A/10T mutant, which activates telomerase but does not target p53 for degradation, is still able to bind the E6-associated protein, E6-AP (J. Huibregtse, personal communication), a ubiquitin-protein ligase necessary for p53 ubiquitination^{19,20}. It is possible that a protein involved in telomerase repression is targeted by E6 for ubiquitin-mediated degradation in a manner similar to that observed for degradation of p53. □

Received 20 November 1995; accepted 15 January 1996.

- Greider, C. W. & Blackburn, E. H. *Cell* **51**, 887–898 (1987).
- Blackburn, E. H. *Rev. Biochem.* **61**, 113–129 (1992).
- Feng, J. L. *et al. Science* **269**, 1236–1241 (1995).
- Morin, G. B. *Cell* **59**, 521–529 (1989).
- Chadeneau, C., Hay, K., Hirte, H. W., Gallinger, S. & Bacchetti, S. *Cancer Res.* **55**, 2533–2536 (1995).
- Counter, C. M. *et al. EMBO J.* **11**, 1921–1929 (1992).
- Counter, C. M., Hirte, H. W., Bacchetti, S. & Harley, C. B. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2900–2904 (1994).
- Kim, N. W. *et al. Science* **266**, 2011–2015 (1994).
- Klingelutz, A. J., Barber, S. A., Smith, P. P., Dyer, K. & McDougall, J. K. *Molec. cell. Biol.* **14**, 961–969 (1994).
- Allsopp, R. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 10114–10118 (1992).
- Hastie, N. D. *et al. Nature* **346**, 866–868 (1990).
- Lindsey, J., McGill, N. I., Lindsey, L. A., Green, D. K. & Cooke, H. J. *Mutat. Res.* **256**, 45–48 (1991).
- Harley, C. B., Futcher, A. B. & Greider, C. W. *Nature* **345**, 458–460 (1990).
- Levy, M. Z., Allsopp, R. C., Futcher, A. B., Greider, C. W. & Harley, C. B. *J. molec. Biol.* **225**, 951–960 (1992).
- Harley, C. B. *Mutat. Res.* **256**, 271–282 (1991).
- Halbert, C. L., Demers, G. W. & Galloway, D. A. *J. Virol.* **65**, 473–478 (1991).
- Miller, A. D., Bender, M. A., Harris, E. A., Kaleko, M. & Gelinas, R. E. *J. Virol.* **62**, 4337–4345 (1988).
- Foster, S. A., Demers, G. W., Etscheid, B. G. & Galloway, D. A. *J. Virol.* **68**, 5698–5705 (1994).
- Huibregtse, J. M., Scheffner, M. & Howley, P. M. *EMBO J.* **10**, 4129–4135 (1991).

20. Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. *Cell* **63**, 1129–1136 (1990).

21. Broccoli, D., Young, J. W. & Delange, T. *Proc. natn. Acad. Sci. U.S.A.* **92**, 9082–9086 (1995).

ACKNOWLEDGEMENTS. We thank D. Galloway for discussion and the retroviral constructs; K. Tedford for technical assistance; D. Hawkins for the human mammary epithelial cells; and T. de Lange for sharing unpublished information on the TRAP assay protocol. Supported by a grant from the NIH.

A mammalian SRB protein associated with an RNA polymerase II holoenzyme

David M. Chao*, Ellen L. Gadbois*, Peter J. Murray*, Stephen F. Anderson†, Michelle S. Sonu*, Jeffrey D. Parvin† & Richard A. Young*

*Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

†Division of Molecular Oncology and Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA

A LARGE multisubunit complex containing RNA polymerase II, general transcription factors and SRB regulatory proteins initiates transcription of class II genes in yeast cells^{1–4}. The SRB proteins are a hallmark of this RNA polymerase II holoenzyme as they are found only in this complex, where they contribute to the response to regulators^{4–8}. We have now isolated a human homologue of the yeast SRB7 gene and used antibodies against human SRB7 protein to purify and characterize a mammalian RNA polymerase II holoenzyme containing the general transcription factors TFIIE and TFIIH. This holoenzyme is more responsive to transcriptional activators than core RNA polymerase II when assayed in the presence of coactivators.

A human complementary DNA clone encoding a protein similar to *Saccharomyces cerevisiae* SRB7 was isolated from a lymphocyte cDNA library by using information derived from expressed sequence tags. The sequence of this human cDNA clone (hSRB7) predicts a 144-amino-acid protein with a relative molecular mass of 15,700 (15.7K) (Fig. 1a). The sequence of the predicted protein is 35% identical to yeast SRB7 (ySRB7) (Fig. 1b). We tested hSRB7's ability to substitute functionally for ySRB7 by determining whether hSRB7 could complement a complete deletion of the yeast gene. Although full-length hSRB7 failed to complement a ySRB7 deletion, several chimaeras containing the amino terminus of the human protein and the carboxy terminus of the yeast protein were able to do so (Fig. 1c, d). The complementing chimaera with the largest amount of hSRB7 contains 57% human sequence.

The yeast SRBs interact functionally and physically with the RNA polymerase II C-terminal repeat domain (CTD; reviewed in ref. 4). If the protein encoded by hSRB7 is a genuine homologue of yeast SRB7, then it should be among a small subset of cellular proteins capable of binding to a recombinant CTD column⁶. Extracts were prepared from yeast, HeLa and calf thymus and

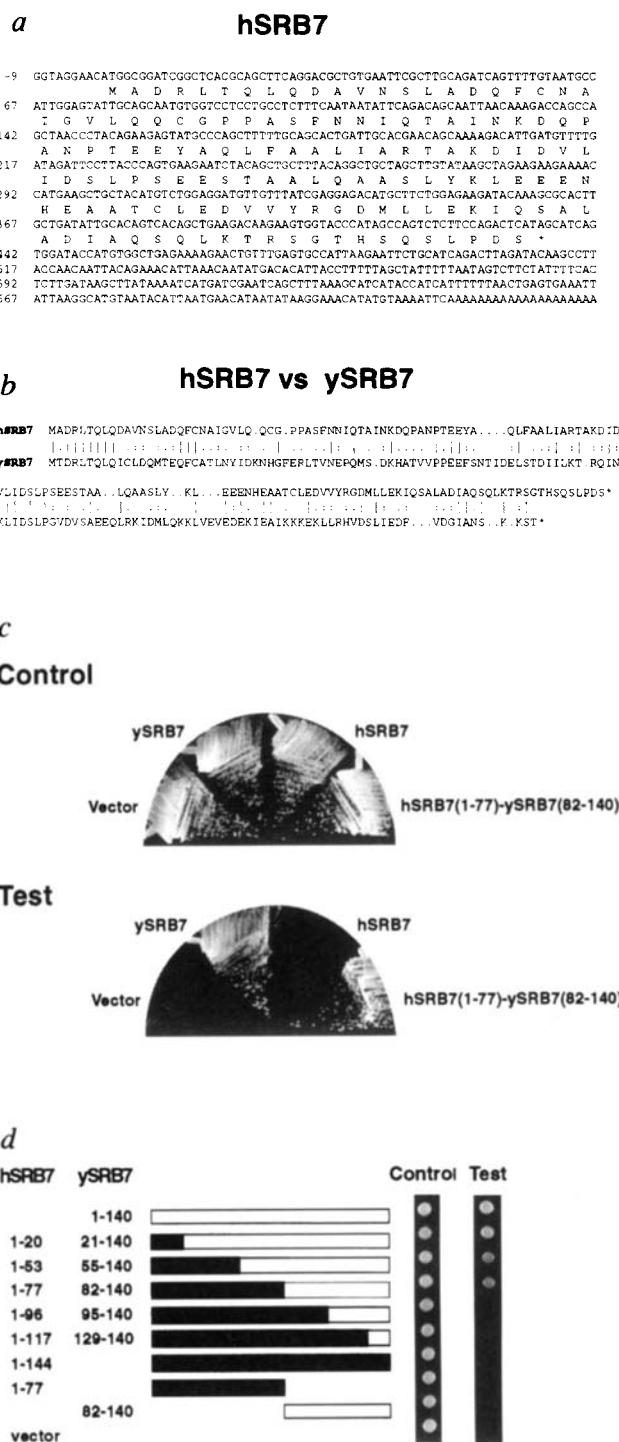


FIG. 1 a, Human SRB7 sequence (GenBank accession number u46837). b, Human and yeast SRB7 alignment. Vertical lines indicate identity; colons, comparison value ≥ 0.5 ; and dots, comparison value ≥ 0.1 as defined by the program BESTFIT. c, Complementation of yeast SRB7

deletion mutant by hSRB7–ySRB7 chimaera. Control: yeast with SRB7 deletion covered by ySRB7 plasmid and the indicated construct. Test: yeast with SRB7 deletion and the indicated construct. d, Complementation by hSRB7–ySRB7 chimaeras. Chimaeras are represented by normalized bars showing human SRB7 sequences in black and yeast sequences in white. METHODS. dbEST was screened with XREFdb for expressed sequence tags similar to ySRB7 (ref. 20). A description of the cloning of hSRB7 and the construction of hSRB7–ySRB7 chimaeras is available as Supplementary Information at Nature's web site (<http://www.nature.com>). ySRB7 and hSRB7 were aligned with BESTFIT (Genetics Computer Group, Inc.). Plasmids containing chimaeras were shuffled into yeast strain Z704 (MATa ura3-52, his3Δ200, leu2-3,112, srb7Δ1 (pCH7: SRB7 URA3 CEN))²¹. A representative clone of each strain was tested by streaking or spotting.

subjected to CTD-affinity chromatography (Fig. 2a). Western blots of the column eluates confirmed that yeast SRB7 was retained on a CTD column and showed that mammalian SRB7 from HeLa cells and calf thymus was also retained (Fig. 2b).

Because yeast RNA polymerase II holoenzyme can be immunoprecipitated with anti-SRB7 antibodies, similar experiments were used to investigate whether mammalian SRB7 associated with components of the transcriptional apparatus in crude extracts. Indeed, RNA polymerase II was specifically immunoprecipitated by anti-hSRB7 antibody (Fig. 2c, d). *In vitro* transcription assays confirmed the presence of RNA polymerase II and revealed the presence of TFIIE and TFIIH activities in the anti-SRB7 immunoprecipitates (Fig. 2e). These results are evidence for a mammalian holoenzyme complex containing at least SRB7, RNA polymerase II, TFIIE and TFIIH.

The complex containing mammalian SRB7 and RNA polymerase II was purified over six columns (Fig. 3a). As shown by western blotting, RNA polymerase II and SRB7 coeluted precisely from the last three columns of the purification. Analysis of material from the last column revealed that SRB7 and RNA polymerase II coeluted with subunits of TFIIE (p56) and TFIIH (p89) (Fig. 3b, c and d). The number of coeluting polypeptides present in the SRB7-RNA polymerase II complex is consistent with the complex's estimated size of M_r 2,000K, as determined by gel filtration chromatography of crude extracts (data not shown). The preparation appears close to purity as defined by coelution of the same set of proteins over two columns. However, *in vitro*

transcription results indicate that TFIIE and TFIIH are substoichiometric (Fig. 4a), suggesting that a portion of their activity was lost by dissociation or inactivation during purification.

We next compared the responses of purified mammalian holoenzyme and core RNA polymerase II to the activator Gal4-VP16 and the coactivators HMG2 (refs 9, 10) and PC4 (refs 11, 12). In four independent experiments, the holoenzyme was more strongly inhibited by PC4 and HMG2 in the absence of activator and showed a modestly enhanced response to activator in the presence of these coactivators (Fig. 4b). Transcription by core RNA polymerase II was mildly inhibited by PC4 and HMG2 (compare lanes 1 and 3) and was stimulated ~2-fold by activator (compare ratio of upper and lower bands in lanes 3 and 4). Transcription by holoenzyme was more strongly inhibited by PC4 and HMG2 (compare lanes 5 and 7) and was stimulated ~5-fold by the activator (compare ratio of upper and lower bands in lanes 7 and 8). It will be interesting to study the holoenzyme's response to other coactivators and cofactors, such as TBP-associated factors (reviewed in ref. 13), topoisomerase I (ref. 14, Dr1/NC2 (refs 15, 16) and PC2 (ref. 17).

We have shown that hSRB7 shares sequence homology with its yeast counterpart, that hSRB7-ySRB7 chimaeras functionally complement a yeast SRB7 deletion, that hSRB7 is specifically retained by a CTD column and, most important, that hSRB7 associates with a transcriptionally active 2,000K complex containing RNA polymerase II and general transcription factors. We conclude that hSRB7 is a genuine homologue of a yeast SRB7 gene

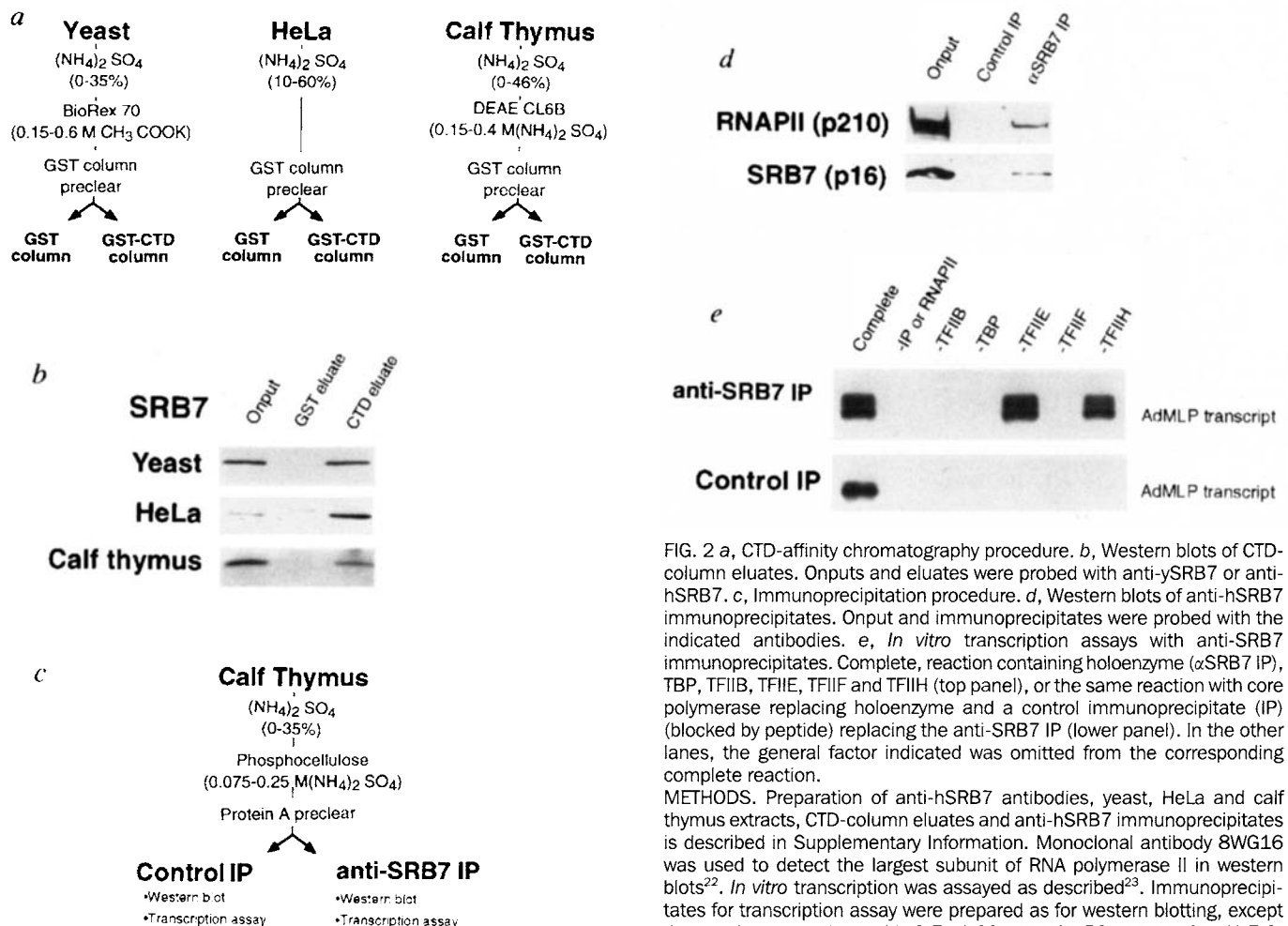


FIG. 2 a, CTD-affinity chromatography procedure. b, Western blots of CTD-column eluates. Onputs and eluates were probed with anti-ySRB7 or anti-hSRB7. c, Immunoprecipitation procedure. d, Western blots of anti-hSRB7 immunoprecipitates. Onput and immunoprecipitates were probed with the indicated antibodies. e, *In vitro* transcription assays with anti-SRB7 immunoprecipitates. Complete, reaction containing holoenzyme (α SRB7 IP), TBP, TFIIB, TFIIE, TFIIIF and TFIIH (top panel), or the same reaction with core polymerase replacing holoenzyme and a control immunoprecipitate (IP) (blocked by peptide) replacing the anti-SRB7 IP (lower panel). In the other lanes, the general factor indicated was omitted from the corresponding complete reaction.

METHODS. Preparation of anti-hSRB7 antibodies, yeast, HeLa and calf thymus extracts, CTD-column eluates and anti-hSRB7 immunoprecipitates is described in Supplementary Information. Monoclonal antibody 8WG16 was used to detect the largest subunit of RNA polymerase II in western blots²². *In vitro* transcription was assayed as described²³. Immunoprecipitates for transcription assay were prepared as for western blotting, except that washes were done with 0.5 ml 60 mM KCl, 50 mM Tris-Cl, pH 7.9, 5 mM MgCl₂, 2.5 mM MnCl₂.

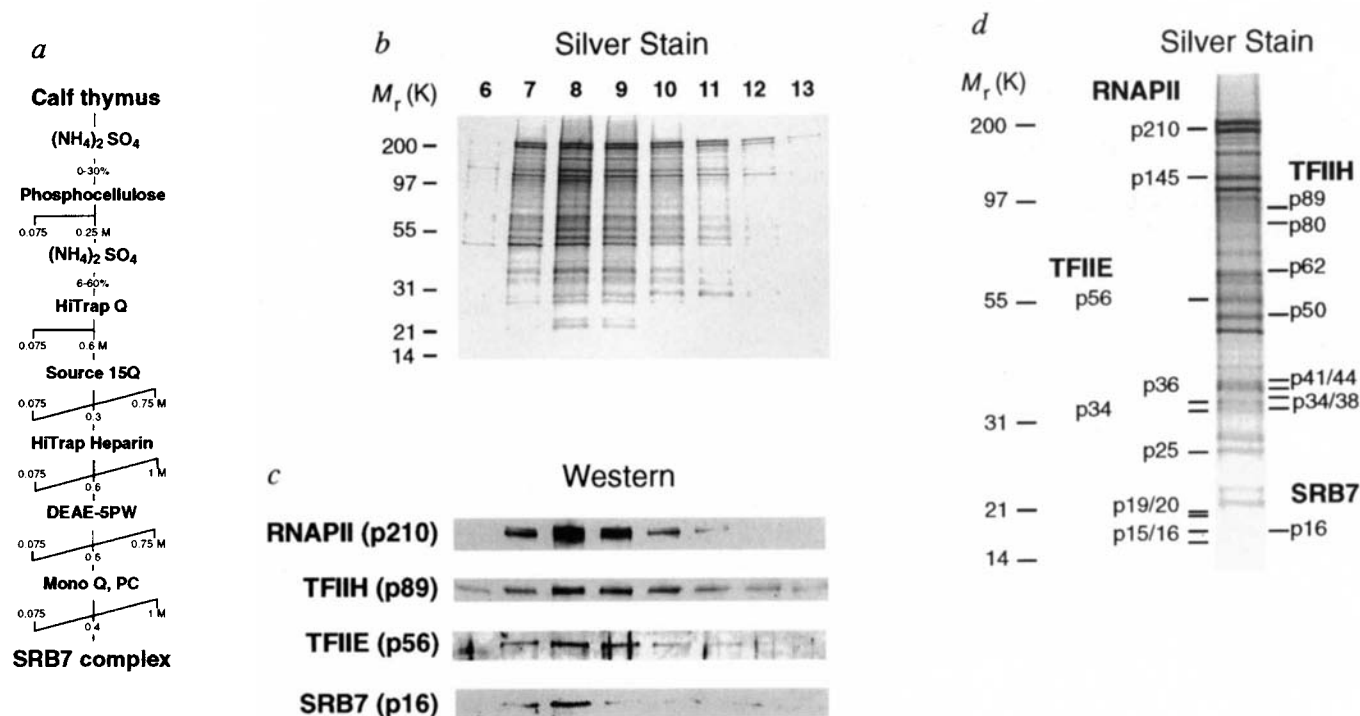


FIG. 3 **a**, Procedure for purifying SRB7. **b** Silver staining of fractions eluted from a Mono-Q column. **c**, Western blotting of fractions eluted from Mono-Q. Fractions were probed with the indicated antibodies. **d**, Proposed identity of holoenzyme polypeptides. Bands that correspond in size to RNA polymerase II, TFIIE and TFIIH subunits and mammalian SRB7 are indicated. METHODS. The purification of mammalian holoenzyme is described in

Supplementary Information. Silver staining of purified holoenzyme has been described⁶. For western blotting, rabbit polyclonal anti-TFIIH p89 and anti-TFIIE p56 (gifts from J. Kim, B. Shykind and P. Sharp) were used at a dilution of 1:500, 3 µl of each fraction was analysed. Identities of holoenzyme polypeptides were assigned based on published compositions of core RNA polymerase II²⁴, TFIIH²⁵, and TFIIE²⁶.

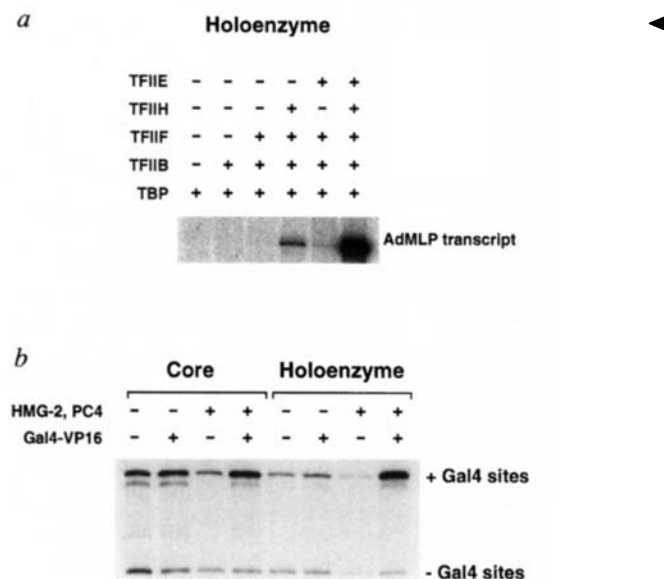


FIG. 4 **a**, *In vitro* transcription assays. Reactions contained column-purified holoenzyme, the general factors indicated, and the adenovirus major late promoter with linear topology. **b**, Response of core RNA polymerase II and column-purified holoenzyme to coactivators and activators. Reactions contained core polymerase or holoenzyme, general transcription factors and/or coactivators and Gal4-VP16. The upper transcript is derived from a template containing the adenovirus major late promoter and three Gal4 binding sites; the lower transcript is derived from a control template containing the same promoter with no Gal4 binding sites. METHODS. Transcription reactions containing TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH have been described²³. Holoenzyme was the peak fraction from the Mono-Q column. Protein preparations for all of the basal factors used here have been shown to be free of cross-contamination²⁷. HMG2 (30 ng per reaction) and PC4 (50 ng per reaction) were titrated for optimal activation.

larger entity. In this context, it is not yet clear whether the yeast holoenzyme contains TBP *in vivo*. Similarly, it remains to be determined whether the *in vivo* form of the mammalian RNA polymerase II holoenzyme contains all¹⁹ or some of the general transcription factors. The isolation of a human SRB gene and a mammalian RNA polymerase II holoenzyme provides a means for investigating these and other issues in transcriptional regulation and extends the holoenzyme model from yeast to mammals. □

Received 3 November 1995; accepted 19 January 1996.

- Koleske, A. J. & Young, R. A. *Nature* **368**, 466-469 (1994).
- Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. E. *Cell* **77**, 599-608 (1994).
- Thompson, C. M. & Young, R. A. *Proc. natn. Acad. Sci. U.S.A.* **92**, 4587-4590 (1995).
- Koleske, A. J. & Young, R. A. *Trends biochem. Sci.* **20**, 113-116 (1995).
- Koleske, A. J., Buratowski, S., Nonet, M. & Young, R. A. *Cell* **69**, 883-894 (1992).
- Thompson, C. M., Koleske, A. J., Chao, D. M. & Young, R. A. *Cell* **73**, 1361-1375 (1993).
- Liao, S. M. *et al.* *Nature* **374**, 193-196 (1995).
- Hengartner, C. J. *et al.* *Genes Dev.* **9**, 897-910 (1995).
- Shykind, B. M., Kim, J. & Sharp, P. A. *Genes Dev.* **9**, 1354-1365 (1995).

and that hRBS7 is a hallmark component of a mammalian RNA polymerase II holoenzyme. We believe that the yeast RNA polymerase II holoenzyme contains RNA polymerase II, SRB proteins and the general factors TFIIB, E, F, and H *in vivo*. Because different holoenzyme purification procedures cause the loss of different subsets of the general transcription factors^{1,2,18}, the forms of holoenzyme purified so far may be subcomplexes of a

10. Stelzer, G., Goppelt, A., Lottspeich, F. & Meisterernst, M. *Molec. cell. Biol.* **14**, 4712–4721 (1994).
11. Ge, H. & Roeder, R. G. *Cell* **78**, 513–523 (1994).
12. Kretzschmar, M., Kaiser, K., Lottspeich, F. & Meisterernst, M. *Cell* **78**, 525–534 (1994).
13. Goodrich, J. A. & Tijan, R. *Curr. Opin. Cell Biol.* **6**, 403–409 (1994).
14. Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J. & Reinberg, D. *Nature* **365**, 227–232 (1993).
15. Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S. & Reinberg, D. *Cell* **70**, 477–489 (1992).
16. Kim, T. K., Zhao, Y., Ge, H., Bernstein, R. & Roeder, R. G. *J. Biol. Chem.* **270**, 10976–10981 (1995).
17. Kretzschmar, M., Stelzer, G., Roeder, R. G. & Meisterernst, M. *Molec. cell. Biol.* **14**, 3927–3937 (1994).
18. Koleske, A. J., Chao, D. M. & Young, R. A. *Meth. Enzym.* (in the press).
19. Ossipow, V., Tassan, J. P., Nigg, E. A. & Schibler, U. *Cell* **83**, 137–146 (1995).
20. Bassett, D. E. J. et al. *Trends Genet.* **11**, 372–373 (1995).
21. Bocke, J. D., Trueheart, J., Natsoulis, G. & Fink, G. R. *Meth. Enzym.* **154**, 164–175 (1987).
22. Thompson, N. E., Aronson, D. B. & Burgess, R. R. *J. Biol. Chem.* **265**, 7069–7077 (1990).
23. Mäkelä, T. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 5174–5178 (1995).
24. Hodo, H. & Blatti, S. P. *Biochemistry* **16**, 2334–2343 (1977).
25. Drapkin, R. & Reinberg, D. *Trends biochem. Sci.* **19**, 504–508 (1994).
26. Ohkuma, Y., Sumimoto, H., Horikoshi, M. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9163–9167 (1990).
27. Parvin, J. D., Shykind, B. M., Meyers, R. E., Kim, J. & Sharp, P. A. *J. Biol. Chem.* **369**, 18414–18421 (1994).

ACKNOWLEDGEMENTS. We thank F. Lewitter for assistance with database searches; S. Elledge for a phage library; L. Guarente for a yeast expression plasmid; L. Strasheim and R. Burgess for an unpublished RNA polymerase II purification protocol; J. Kim, B. Shykind and P. Sharp for antibodies and purified transcription factors; and G. Fink, P. Sharp, J. and R. Conaway, D. Reinberg, R. Roeder and U. Schibler for advice and discussion. J.D.P. thanks R. S. Cotran for his support. D.M.C. and E.L.G. are predoctoral fellows of the Howard Hughes Medical Institute. S.F.A. was supported by an institutional training grant from the NIH. This work was supported by NIH grants to R.A.Y.

Initiation of V(D)J recombination *in vitro* obeying the 12/23 rule

Quinn M. Eastman*, Thomas M. J. Leu† & David G. Schatz‡

Howard Hughes Medical Institute, Section of Immunobiology,

* Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, 310 Cedar Street, PO Box 208011, New Haven, Connecticut 06520–8011, USA

† Universität Zürich-Irchel, Veterinärbiochemie, Winterthurerstrasse 190, 8057 Zürich, Switzerland

V(D)J RECOMBINATION, the process that assembles antigen-receptor genes, is directed by signal sequences flanking the DNA segments to be joined. Signals consist of a conserved heptamer and nonamer separated by a spacer of either 12 or 23 base pairs. Recombination occurs almost exclusively between two signals with spacers of different lengths. This restriction, called the '12/23 rule', governs the organization and pattern of rearrangement of antigen-receptor loci^{1,2}. *In vitro* work demonstrating the direct roles of the Rag proteins in the initiation of V(D)J recombination did not recreate the 12/23 rule^{3,4}. Instead, double-strand breaks were formed efficiently at isolated signals. Here we show that extracts made from a lymphoid cell line that expresses truncated forms of the Rag1 and Rag2 proteins have a signal-cutting activity that obeys the 12/23 rule. Cleavage at the two signals is concerted and requires their synapsis, and mutations of one signal prevent cleavage at both.

The prototype substrate, p12 × 23 (Fig. 1a), is a 531-base-pair (bp) product of the polymerase chain reaction (PCR), uniformly labelled with ³²P. It contains a recombination signal with a 12-bp spacer (12-signal) and another with a 23-bp spacer (23-signal). The other substrates used in the initial assays have different combinations or orientations of 12- and 23-signals (Fig. 1b). DNA cleavage during V(D)J recombination occurs at the heptamer border of the signals and can be mimicked by cutting with *Sal*I and/or *Bam*HI (Fig. 1c, lanes 2–4). When incubated with extract containing the Rag1 and Rag2 proteins, p12 × 23 was cut at both

signals, generating 279-, 175- and 77-bp fragments (lane 7). The latter is visible on longer exposures as a broad smear. The smearing of the 175- and 77-bp fragments is probably a result of action by nonspecific nucleases in the extracts.

The presence of both a 12-signal and a 23-signal are critical for cleavage. No cleavage is detected at the isolated signals of p12 and p23 (Fig. 1c, lanes 5 and 6). p23 × 23 is also not cleaved detectably (lane 9), whereas cleavage of p12 × 12 is barely detectable (lane 8), demonstrating that a 12-signal and a 23-signal cannot efficiently substitute for one another. Detectable cleavage of p12 × 12 is consistent with the observation that substrates containing two similar signals recombine at low levels *in vivo*^{5,6}. Furthermore, there is no visible cleavage in a reaction containing both p12 and p23, suggesting that the two signals must be present on the same DNA molecule (lane 10). We estimate that bands resulting from cleavage of less than ~1% of total substrate would be obscured by background. pD243 and pC329, which contain both a 12-signal and a 23-signal, oriented differently from those in p12 × 23 (Fig. 1b), are cleaved with an efficiency comparable to p12 × 23 (Fig. 1c, lane 11, and Fig. 3c, lane 2). Our results are in good agreement with data from cultured cells showing that efficient cleavage of V(D)J recombination substrates *in vivo* requires a 12- and a 23-signal (D. B. Roth, personal communication). Therefore, within the limits of sensitivity of our assay, the *in vitro* cleavage reaction conforms to the 12/23 rule and recapitulates an important aspect of the regulation of V(D)J recombination *in vivo*. These results also demonstrate that the large portions of the Rag proteins deleted in the truncation mutants used here are not required for 12/23-regulated cleavage.

Extracts from cells expressing only Rag1 or only Rag2 have no visible activity (Fig. 1d, lanes 4 and 5). A divalent cation is required: when Mg²⁺ is replaced by 0.5 mM EDTA, there is no visible cleavage activity (lane 6). Consistent with previous studies^{7,8}, the majority of coding ends formed in our reactions are covalently sealed hairpins and the signal ends are not, as determined by analysis of cleavage products under denaturing conditions (data not shown).

We have not observed cleavage at only one signal in standard cleavage reactions using p12 × 23 (Fig. 1c, lane 7 and d, lanes 1–3), pD243 (Fig. 1c, lane 11), or pC329 (Fig. 3c, lane 2). To look for singly cut intermediates, we performed cleavage reactions with p12 × 23 at several temperatures (25, 30 and 37 °C), and removed aliquots at early time points. No singly cleaved DNA fragments were visible at any time or at any temperature tested (data collected at 30 °C are shown in Fig. 2). These results support a model in which there is tight temporal coupling between cleavage events at the two signals, for which there is some support from studies of V(D)J recombination *in vivo*⁹. This is in contrast to asynchronous cleavage seen at the ends of the bacterial transposons Tn7 and Tn10 (refs. 10,11) and the asynchronous processing of retroviral ends^{12–14}.

Previously reported *in vitro* cleavage studies primarily used Mn²⁺ and found cleavage at isolated 12- or 23-signals^{3,4}, whereas our standard reaction contains 10 mM Mg²⁺. When we substituted Mn²⁺ for Mg²⁺, we observed a reduction (~10-fold) in the efficiency of cleavage and an alteration in the specificity of the cleavage reaction in two regards: first, in the presence of Mn²⁺, p12 × 23 undergoes single cleavage at the 12-signal, and second, p12 × 12 is now cleaved more efficiently, with cleavage occurring at the equivalent 12-signal (Fig. 1d, lanes 8–10, 12–14; arrow indicates the singly cleaved product). Mn²⁺ is known to stimulate misincorporation by DNA polymerases and relax the specificity of some restriction endonucleases^{15,16}. Mn²⁺ may act in a similar way here, stabilizing otherwise unfavourable interactions between the Rag proteins and the DNA.

The requirement for both a 12-signal and a 23-signal raised the possibility that synapsis of the two signals was a prerequisite for cleavage. To test this idea we performed cleavage reactions with a series of substrates that systematically reduced the distance between the signals, raising the energetic cost of the DNA bending

‡To whom correspondence should be addressed.

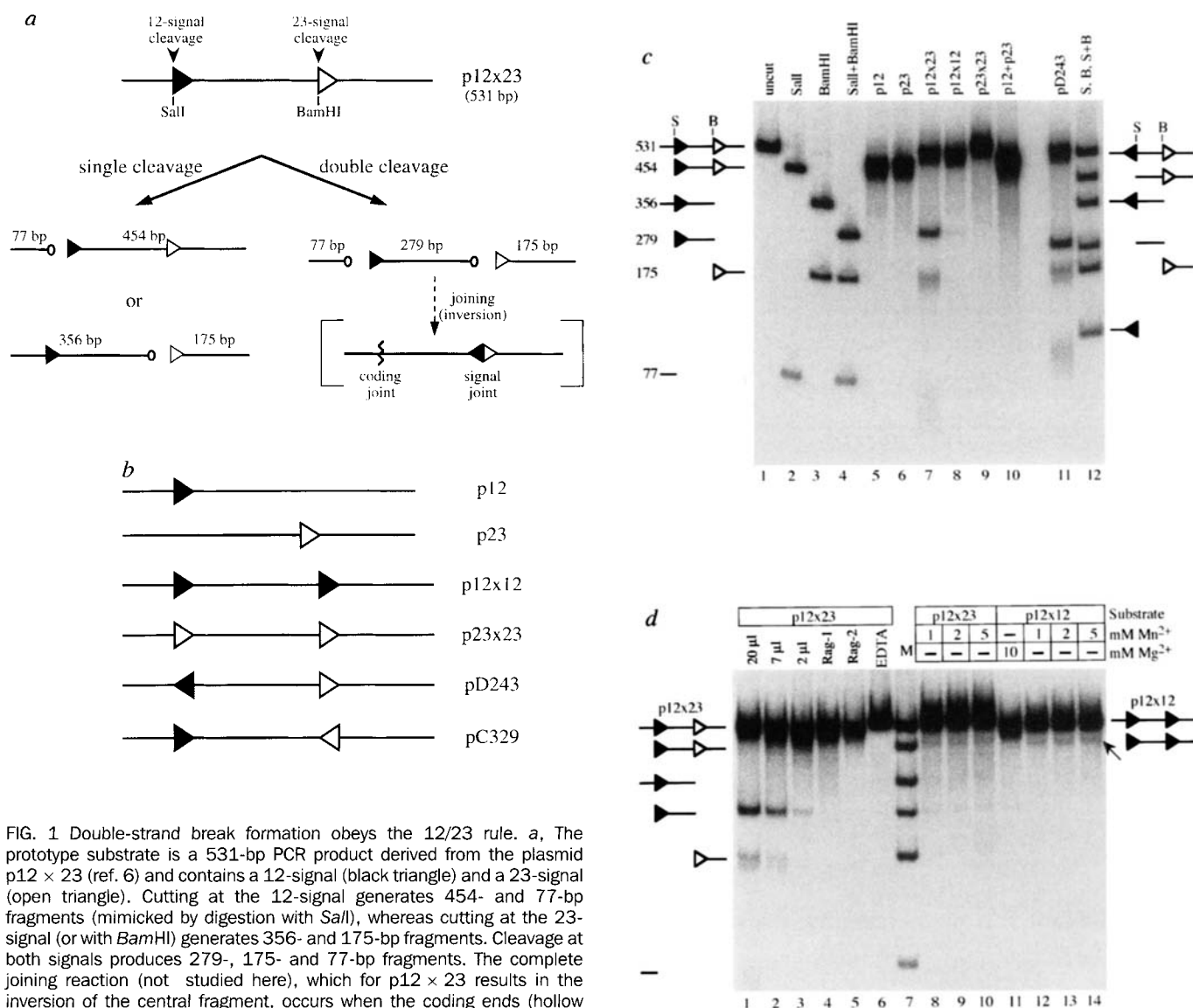
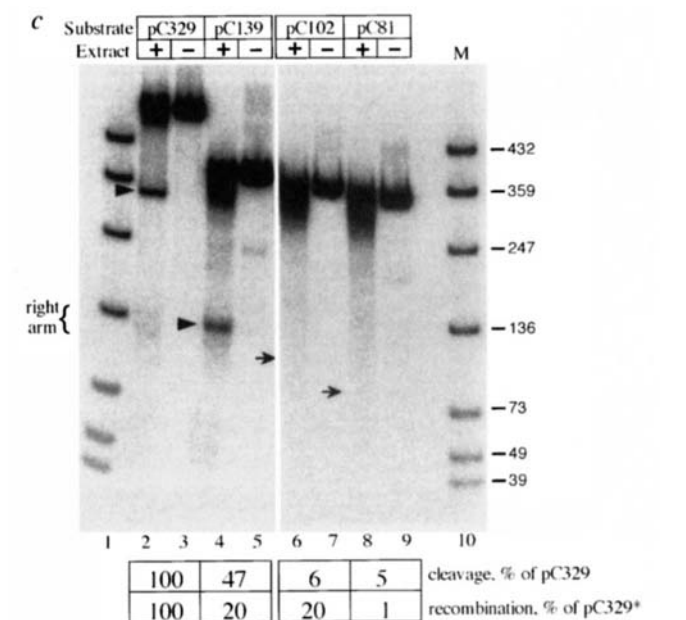
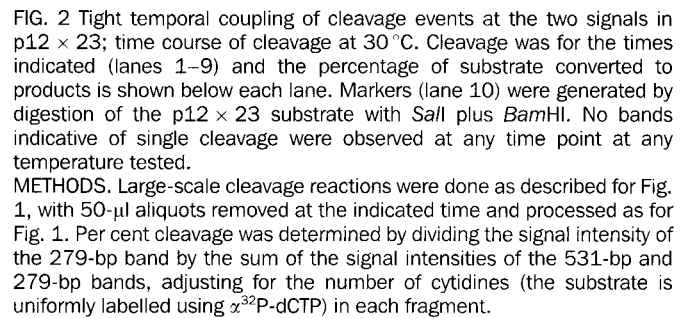


FIG. 1 Double-strand break formation obeys the 12/23 rule. **a**, The prototype substrate is a 531-bp PCR product derived from the plasmid p12 × 23 (ref. 6) and contains a 12-signal (black triangle) and a 23-signal (open triangle). Cutting at the 12-signal generates 454- and 77-bp fragments (mimicked by digestion with *Sall*), whereas cutting at the 23-signal (or with *Bam*HI) generates 356- and 175-bp fragments. Cleavage at both signals produces 279-, 175- and 77-bp fragments. The complete joining reaction (not studied here), which for p12 × 23 results in the inversion of the central fragment, occurs when the coding ends (hollow ovals) are ligated together to form a coding joint (zigzag line), and the two signals are ligated together to form a signal joint. **b**, Substrates containing a single signal (p12, p23), two similar signals (p12 × 12 and p23 × 23), or two different signals, with the signals in different relative orientations compared with those in p12 × 23 (pD243 and pC329). In pD243 and pC329, the complete recombination reaction results in deletion of the central fragment, with retention of a signal joint or a coding joint, respectively, on the remaining DNA. **c**, Cleavage reactions demonstrating the 12/23 rule *in vitro*. Reactions were performed with the indicated substrate or combination of substrates. Markers were obtained by restriction digestion of the p12 × 23 substrate (lanes 2–4; lane 1 shows the undigested substrate, and the size of markers in bp is shown on the left) or the pD243 substrate (a mixture of uncut, *Sall*-digested, *Bam*HI-digested, and *Sall*-plus-*Bam*HI-digested; lane 12). The structure of each cleavage product is indicated on the left for p12 × 23 and on the right for pD243. Cleavage efficiencies in this experiment were: p12 × 23, 26.0%; pD243, 32.5%; p12 × 12, 1%. **d**, Requirements for cleavage and the 12/23 rule. Cleavage reactions were performed with p12 × 23 and decreasing amounts of the complete extract (lanes 1–3), or extracts containing only Rag1 (lane 4), or only Rag2 (lane 5), or the complete extract with 0.5 mM EDTA in place of Mg²⁺ (lane 6). Mn²⁺ (at the indicated concentrations) was substituted for Mg²⁺ using either p12 × 23 (lanes 8–10) or p12 × 12 (lanes 12–14) as the substrate. Under these conditions the substrates were cut predominantly at the leftmost 12-signal (band indicated by the arrow, representing 2–3% or 3–5% of the total substrate for p12 × 23 and p12 × 12, respectively). For comparison, the standard reaction (10 mM Mg²⁺) with p12 × 12 is shown in lane 11. In this experiment, 18.3% of p12 × 23 was cleaved (lane 1). The pattern of bands generated from p12 × 23 with 1–5 mM Mg²⁺ is the same as with 10 mM, but cleavage is less efficient. Markers (lane M) are of a mixture of the samples from lanes 1–4 in **c**.

METHODS. M12 cells were stably transfected with truncated RAG1 (pMS127; ref. 18) and RAG2 (R2CC14; ref. 19) cDNAs under heat-shock control²⁰. Frozen pellets of induced cells were extracted with buffer containing 300 mM NaCl. Proteins were precipitated with 75% saturated (NH₄)₂SO₄ and dialysed against 50 mM NaCl, 25 mM Tris-HCl, pH 8.0, 50 µM ZnSO₄, 2 mM DTT, 20% glycerol. Extracts contained ~12 mg ml⁻¹ total protein, and 50 µg ml⁻¹ Rag1 and 2 µg ml⁻¹ Rag2, as judged by semi-quantitative western blotting with Rag-specific antibodies²⁰. M12 lines expressing Rag1 or Rag2 only contained similar levels of Rag proteins. Labelled substrates were generated in PCR reactions (16 cycles) containing 50 µCi α³²P-dCTP, 50 µM dNTPs and 16 ng linear plasmid template, and were purified on Qiaquick columns (QIAGEN). PCR primers: LE1, 5'-GGAATTGTGAGCGGATAAC-3'; cit4a, 5'-GAGGCATTTTCAGTCAGTGC-3'. Plasmid templates have been described⁶, except for p12 × 12, which was created by ligating the annealed oligonucleotides 5'-GATCCACAGT-GATTGATATCACTACAAAAACC-3' and 5'-GATCGGTTTTGTAGTGATATCAAT-CACTGTG-3' into the *Bam*HI site of p12. A standard 50-µl cleavage reaction contained 1.2–1.5 × 10⁵ c.p.m. substrate (10 ng DNA), 20 µl extract from cells expressing both Rag1 and Rag2, 150 mM sodium acetate, 10 mM magnesium acetate, 25 mM Na-HEPES, pH 7.5, 1 mM EGTA, and was incubated at 37 ° for 2 h. Reactions were stopped by adding 0.5% SDS and excess EDTA, digested with proteinase K, alcohol-precipitated, resuspended and loaded on a 4% non-denaturing polyacrylamide 1 × TBE gel. This and other figures were generated from scans of the dried gels using a BioRad Molecular Imager and Molecular Analyst software.



87

required for synapsis. Previous studies have shown that short intersignal distances substantially inhibit V(D)J recombination *in vivo*^{6,17}. For inversion and coding joint substrates, reducing the intersignal distance (as defined in Fig. 3a) to 105 bp or 102 bp, respectively, virtually abolished cleavage, whereas for signal joint substrates, cleavage was reduced fivefold, with an intersignal distance of 40 bp (Fig. 3b, c). These data are in good general agreement with *in vivo* recombination results¹⁷ with the same substrates (see comparison at the bottom of Fig. 3b, c).

The sequence requirements for V(D)J recombination have been studied by extensive mutational analysis of the signals in plasmid recombination substrates⁵. To test whether our signal-cutting activity had a similar specificity and to determine whether

mutations at one signal would affect the efficiency of cleavage at both, we performed cleavage reactions with a panel of mutant recombination substrates. The mutants studied here contain one-base changes in the heptamer or the nonamer of the 12-signal or the nonamer of the 23-signal, or changes in the lengths of one of the two spacers (Fig. 4a). The wild-type substrate, pJH299, is similar to p12 × 23 in signal orientation and is cleaved efficiently (Fig. 4b, lane 1). With the mutant substrates (Fig. 4b, lanes 2–10), the efficiency of cleavage correlates well with the efficiency of recombination of the corresponding plasmid *in vivo*⁵; mutations that have a severe effect upon recombination significantly reduce the efficiency of cleavage (Fig. 4b, bottom). For the mutants that are cleaved detectably, the pattern of bands generated by cleavage remains the same as for pJH299. The data emphasize that the recombination machinery must interact efficiently with both signals for cleavage at either signal to occur. An apparent exception is a mutation in the first base of the heptamer of the 12-signal that eliminates detectable cutting at the 12-signal, but allows some cutting at the 23-signal (indicated by the arrow in Fig. 4b, lane 2). A possible explanation is that this mutant 12-signal retains some ability to be bound by the relevant factor(s) and thus allows a degree of complex formation, but disrupts cleavage at the 12-signal.

Recent studies using pre-B cell extracts supplemented with recombinant Rag1 produced in baculovirus-infected insect cells or purified recombinant Rag1 and Rag2, have demonstrated the direct roles played by the Rag proteins in double-strand break formation^{3,4}. In striking contrast with the results shown here, these studies found that cleavage occurred as efficiently at an isolated signal as at a 12-signal and 23-signal together, and was substantially more efficient with Mn²⁺ than with Mg²⁺. Here, cleavage is dependent upon interaction between the 12-signal and the 23-signal. What allows formation of a higher-order complex that brings the two signals together and prevents cleavage at an isolated signal, and what accounts for the absence of proper regulation in previous studies? As shown here, the choice of divalent cation has a dramatic effect in our reactions, and substitution of Mg²⁺ for Mn²⁺ may help explain the discrepancy. In addition, co-expression of the Rag proteins in a mammalian context (unlike the previous studies) might allow proper post-translational modification or aid in higher-order protein complex assembly. It is also possible that other proteins present in our extracts may act to regulate the Rag proteins, restricting efficient cleavage to substrates containing a 12-signal and a 23-signal. Reconstitution of properly regulated signal-cutting activity from purified components, followed by detailed mechanistic studies, should provide an explanation of the 12/23 rule at the molecular level.

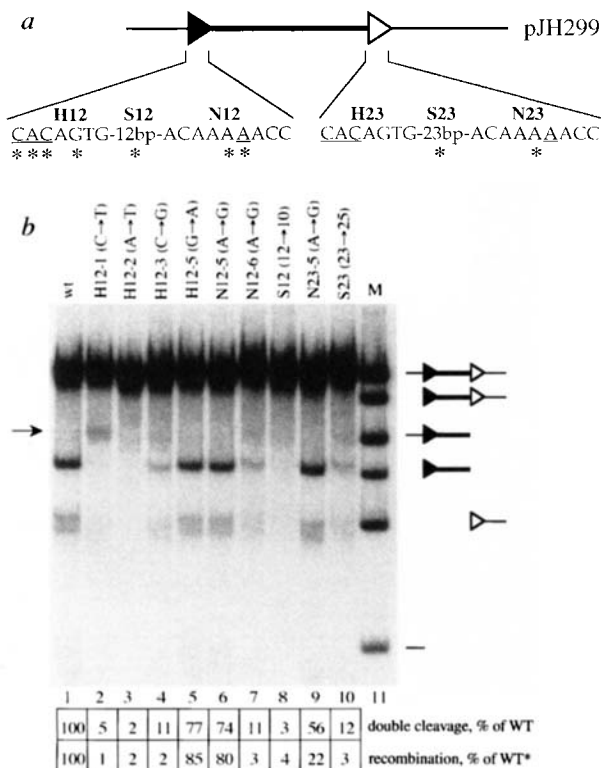


FIG. 4 Mutations of one signal prevent cleavage at both signals. *a*, Configuration and nucleotide sequence (5' to 3') of the signal sequences of the wild-type substrate pJH299. The heptamer, spacer and nonamer of the 12-signal are labelled H12, S12 and N12, respectively, with a similar nomenclature used for the elements of the 23-signal. Each heptamer or nonamer mutant substrate contained a single base change relative to wild type, with the positions of the mutations indicated by asterisks. Spacer mutants contained a change of 2 bp in the length of one or the other spacer. The most highly conserved residues of the heptamer and nonamer are underlined (> 95% identical in a compilation of the sequences of 226 signals⁵). *b*, Cleavage reactions with wild type and mutant substrates. The mutation is described above each lane: H12-*n* refers to the *n*th position of the heptamer in the 12-signal (counting from left to right as in *a*), N12-*i* to the *i*th position of the nonamer in the 12-signal, and N23-5 to the fifth position of the nonamer of the 23-signal. At the bottom is a comparison of the *in vitro* cleavage efficiency for each substrate and the *in vivo* recombination efficiency observed with these substrates in a previous study⁵. The per cent cleavage and per cent recombination are calculated relative to the values obtained for pJH299 (lane 1), which is arbitrarily assigned a value of 100%. The most prominent product band in lane 2 corresponds to the larger fragment resulting from cleavage at the 23-signal only, and represents 6.6% of the substrate in that lane. By comparison, 16.2% of pJH299 was doubly cleaved in lane 1. Lane 11 is a marker lane identical to lane 7 in Fig. 1d. * *In vivo* recombination data were derived from ref. 5. **METHODS.** pJH299 and the mutant substrates were supplied by J. Hesse and M. Gellert⁵. Cleavage reactions were processed and quantified as for Figs 1 and 2.

Received 21 December 1995; accepted 15 February 1996.

1. Tonegawa, S. *Nature* **302**, 575–581 (1983).
2. Lewis, S. M. *Adv. Immun.* **56**, 27–150 (1994).
3. van Gent, D. C. et al. *Cell* **81**, 925–934 (1995).
4. McBlane, J. F. et al. *Cell* **83**, 387–395 (1995).
5. Hesse, J. E., Lieber, M. R., Mizuuchi, K. & Gellert, M. *Genes Dev.* **3**, 1053–1061 (1989).
6. Lewis, S. M. & Hesse, J. E. *EMBO J.* **10**, 3631–3639 (1991).
7. Roth, D. B., Menetsky, J. P., Nakajima, P. B., Bosma, M. J. & Gellert, M. *Cell* **70**, 983–991 (1992).
8. Zhu, C. M. & Roth, D. B. *Immunity* **2**, 101–112 (1995).
9. Roth, D. B., Nakajima, P. B., Menetsky, J. P., Bosma, M. J. & Gellert, M. *Cell* **69**, 41–53 (1992).
10. Bainton, R., Gamas, P. & Craig, N. L. *Cell* **65**, 805–816 (1991).
11. Haniford, D. B. & Kleckner, N. *EMBO J.* **13**, 3401–3411 (1994).
12. Lee, Y. M. H. & Coffin, J. M. *Molec. cell. Biol.* **11**, 1419–1430 (1991).
13. Kulkosky, J., Katz, R. A. & Skalka, A. M. *J. acqu. Immun. Def. Syndr.* **3**, 852–858 (1990).
14. Vora, A. C., Fitzgerald, M. L. & Grandegennett, D. P. *J. Virol.* **65**, 5656–5659 (1990).
15. Copeland, W. C., Lam, N. K. & Wang, T. S. F. *J. Biol. Chem.* **268**, 11041–11049 (1993).
16. Vermote, C. L. M. & Halford, S. E. *Biochemistry* **31**, 6082–6089 (1992).
17. Sheehan, K. M. & Lieber, M. R. *Molec. cell. Biol.* **13**, 1363–1370 (1993).
18. Sadofsky, M. J., Hesse, J. E., McBlane, J. F. & Gellert, M. *Nucleic Acids Res.* **22**, 5644–5650 (1994).
19. Cuomo, C. A. & Oettinger, M. A. *Nucleic Acids Res.* **22**, 1810–1814 (1994).
20. Leu, T. M. J. & Schatz, D. G. *Molec. cell. Biol.* **15**, 5657–5670 (1995).

ACKNOWLEDGEMENTS. We thank M. Lieber, S. Lewis, J. Hesse, M. Sadofsky, M. Oettinger and M. Gellert for plasmids; C. M. Thompson for important suggestions concerning the cleavage assay; and N. Grindley, S. Lewis and C. McMahon for critically reading the manuscript. This work was supported by grants to D.G.S. from the NIH. Q.M.E. was supported by a predoctoral fellowship from the NSF. T.M.J.L. was supported by a grant administered by the Union Bank of Switzerland. D.G.S. is an Assistant Investigator of the Howard Hughes Medical Institute.